Efficient Implementation of the Heston-Hull&White Model

Sheldon Maze*

21st May 2014

Supervisors
Moses dos Santos
Marchand van Rooyen

A dissertation submitted to the Department of Actuarial Science, Faculty of Commerce, University of Cape Town, in partial fulfilment of the requirements for the Degree of
Master of Philosophy
in
Mathematical Finance.

University of Cape Town,
South Africa

*E-mail address: mzxshe001@myuct.ac.za or sheldon.maze@gmail.com.

Abstract

A model with a stochastic interest rate process correlated to a stochastic volatility process is needed to accurately price long-dated contingent claims. Such a model should also price claims efficiently in order to allow for fast calibration. This dissertation explores the approximations for the characteristic function of the Heston-Hull&White model introduced by Grzelak and Oosterlee (2011). Fourier-Cosine expansion pricing, due to Fang and Oosterlee (2008), is then used to price contingent claims under this model, which is implemented in MATLAB. We find that the model is efficient, accurate and has a relatively simple calibration procedure. In back-tests, it is determined that the Heston-Hull&White model produces better hedging profit and loss results than a Heston (1993) or a Black and Scholes (1973) model.

Keywords: Heston Hull & White, stochastic interest rate, stochastic volatility, Fourier-Cosine pricing.

Contents

List of Figures

List of Tables

1 Introduction

The Heston-Hull&White (HHW) model combines stochastic volatility and stochastic interest rates, as described by Grzelak and Oosterlee (2011). This allows more accurate pricing of hybrid derivatives; i.e. derivatives whose payoff is linked to both equity returns and interest rates, a simple example whould be a bet on the correlation between the two (Hunter, 2005). The HHW model allows interest rate and equity returns to be correlated, thus allowing us to price correlation bets, while still being tractable and efficient.

The aim of this dissertation is to implement a version of the model described by Grzelak and Oosterlee (2011) in such a way that it efficiently prices contingent claims. This involves coding the model, calibrating the model and conducting a benchmarking and hedging analysis of the model.

The model will be calibrated to current market data. This requires the model to be run many times, and thus, an efficient implementation of the model is required. If the model is inefficient (i.e. too slow) the calibration process will take too long, which is disadvantageous for any model. Numerical assessments of the model's efficiency, accuracy and hedging ability will also be conducted.

In the following two chapters, we introduce the models discussed in this dissertation and the approximations made by Grzelak and Oosterlee (2011). Then, in Section 4, we explore an efficient implementation in MATLAB using the Fourier-Cosine method. In Section 5, we look at a simple calibration procedure. The models hedging ability is explored and tested in Section 6. Finally, we summarise the findings, and draw conclusions in Section 7.

2 The Models

This section will introduce the models discussed in this dissertation.

2.1 Heston

The stochastic volatility Heston (1993) model is described by the following Stochastic Differential Equations (SDEs):

$$\begin{aligned} dS_t = &\quad S_t r_t dt + \quad S_t\sqrt{v_t}dW_t^x, \\ dv_t = &\quad \kappa(\bar{v} - v_t)dt + \quad \gamma\sqrt{v_t}dW_t^v, \quad v_0 > 0. \end{aligned} \tag{1}$$

with $r > 0$ the constant interest rate and W_t^x, W_t^v two correlated Brownian Motions, with $dW_t^x dW_t^v = \rho_{x,v}dt$ where $\rho_{x,v}$ is the correlation between the two Brownian Motions. The square-root volatility process, v_t, is mean reverting at a rate of $\kappa > 0$ to its long-run mean, $\bar{v} > 0$. The volatility of the volatility process is specified by $\gamma > 0$, and, S_t is the stock price process.

An analytical solution for this model can be found, but, numerical methods such as integral calculation or transform techniques are required to evaluate it. We can, however, use its characteristic function (ChF) to price options.[1] The Heston model is not in an affine form as it is presented in (1), but after applying a log transform to the stock, $x_t = \ln(S_t)$, the model is affine (Heston, 1993).[2] Applying Ito's Lemma and this transformation to (1) results in:

$$\begin{aligned} dx_t = &\quad \left(r_t - \frac{1}{2}v_t\right)dt + \quad \sqrt{v_t}dW_t^x, \\ dv_t = &\quad \kappa(\bar{v} - v_t)dt + \quad \gamma\sqrt{v_t}dW_t^v, \quad v_0 > 0. \end{aligned} \tag{2}$$

The reason for including stochastic volatility in a market model is to account for the volatility skew or smile that is observed in practice. The Black and Scholes (1973) option pricing model assumes a constant volatility and is therefore unable to account for this skew.

[1] See Appendix A for a definition of a ChF.

[2] See Appendix A for a definition of affinity.

2.2 Hull&White

In practice, interest rates are not constant, nor are they deterministic: we therefore introduce a stochastic interest rate model. The Hull&White model is based on a Generalized Ornstein-Uhlenbeck process, with a time dependent, deterministic long-run average θ_t (Hull and White, 1990). The SDE for the Hull&White model is:

$$dr_t = \quad \lambda(\theta_t - r_t)dt + \quad \eta dW_t^r, \quad r_0 > 0, \tag{3}$$

where λ is the rate of mean reversion, θ_t is chosen so that the model exactly fits the interest rate term structure currently being observed, η is the volatility of the interest rate and W_t^r is a standard Brownian Motion.

When pricing long-dated options (i.e. with a maturity greater than one year), the assumption of a constant interest rate causes mispricings. The magnitude and direction (over or under) of these mispricings are dependent on the differences between the constant interest rate and the properties of the stochastic interest rate (e.g. correlation with the underlying), as well as the time to maturity of the option.

The longer the time to maturity, the more time the stochastic interest rate has to reach its long-run average, and the more time this average has to affect the option price. The rate of mean reversion and the volatility of the interest rate process also play a role in how soon the long-run interest rate is reached and how stable it is. If, for example, the long-run average interest rate is above the constant interest rate, assuming a constant rate will under-price 10+ year call options and 20+ year put options.[3]

It has already been noted that, in practice, interest rates are not constant, nor are they deterministic. Thus, adding a stochastic interest rate to a model allows for the correct discounting of future cash flows (Bakshi

[3]This was discovered by comparing vanilla European option prices under the Black-Scholes, Heston and HHW models.

et al., 2000). We would therefore like to include stochastic interest rates in a model pricing long-dated options (Brigo and Mercurio, 2007).

2.3 Heston-Hull&White

When pricing hybrid contingent claims, we need to model interest rates and the stock price process.

Bakshi *et al.* (2000) conducted an analysis on pricing and hedging long-dated option contracts. They specifically looked at hedging a 2-3 year LEAPS (Long-term Equity Anticipation Security) put option on the S&P500. It was found that a medium-term (0.5 to 1 year) put option is the best hedging instrument (compared to the underlying and a short-term option), as it produced the lowest hedging profit and loss and volatility in most cases (Bakshi *et al.*, 2000).[4]

Bakshi *et al.* (2000) compared stochastic volatility; stochastic volatility with jumps; and stochastic volatility with stochastic interest rate models in their tests. They concluded that in most cases, the best model for hedging a 2-3 year LEAPS put option was a stochastic volatility and stochastic interest rate model. It is also noted by Bakshi *et al.* (2000) that the Black-Scholes model had a significantly worse hedging performance than models incorporating stochastic volatility when hedging this option.

These are promising results, as the HHW model also incorporates stochastic volatility and stochastic interest rates. However, in the study conducted by Bakshi *et al.* (2000), the interest rate process was not correlated to any other processes. Hunter (2005) explains that the simplest form of a hybrid derivative is a bet on the correlation between two asset classes. If the hybrid involves equity and cash, bonds or any other product dependent on interest rates, then the hybrid is sensitive to correlations between these products.

[4]Options with moneyness from 0.94 to 1.06 were considered as well as different hedging periods and models. See Bakshi *et al.* (2000) for details.

Thus, the equity and interest rate processes need to be correlated to accurately price these hybrid contingent claims (Hunter, 2005).

Zhu (2000) presents another method to combine stochastic volatility and stochastic interest rate models. However, this method is also unable to have correlated interest rate and equity processes.

Thus, Grzelak and Oosterlee (2011) have combined the *Heston* and *Hull & White* models in such a way that allows this correlation to be non-zero. The HHW model is described by the following SDEs:

$$\begin{aligned} dx_t = &\left(r_t - \frac{1}{2}v_t\right)dt + \sqrt{v_t}dW_t^x, \quad x_t = \ln(S_t), S_0 > 0, \\ dv_t = &\kappa(\bar{v} - v_t)dt + \gamma\sqrt{v_t}dW_t^v, \qquad v_0 > 0, \\ dr_t = &\lambda(\theta_t - r_t)dt + \eta dW_t^r, \qquad r_0 > 0, \end{aligned} \tag{4}$$

where $dW_t^x dW_t^r = \rho_{x,r}dt$, $dW_t^r dW_t^v = \rho_{r,v}dt$ and all other parameters are as before.

As with the Heston model, if we used S_t, the model would not be affine. However, even after the log transformation, the HHW model is not in affine form. Therefore, Grzelak and Oosterlee (2011) provide approximations for the non-affine component of this model.

3 Approximations

The HHW model (as presented above) is not affine. Below we will present the approximations that Grzelak and Oosterlee (2011) produced to allow the ChF to be solved analytically.

For affinity to hold, the expectations and the covariance matrix of the system need to be linear in the state space variables. In (4) x_t, v_t and r_t are the state space variables, which form the following state space vector: $\underline{s_t} = [x_t, v_t, r_t]$. Then, the system (4) can be rewritten as:

$$d\underline{s_t} = \underline{\mu}\left(t, \underline{s_t}\right) dt + \sqrt{\mathbf{\Sigma}} d\underline{W_t}, \tag{5}$$

where the square-root of a matrix is defined such that $\sqrt{\mathbf{A}}\left(\sqrt{\mathbf{A}}\right)^{\mathsf{T}} = \mathbf{A}$. It is clear that the expectations will be linear, as the drifts in (4) are linear and the expectation of Brownian Motion is zero. The instantaneous covariance matrix for the system (4) is:

$$\mathbf{\Sigma} = \begin{bmatrix} v_t & \rho_{x,v}\gamma v_t & \rho_{x,r}\eta\sqrt{v_t} \\ * & \gamma^2 v_t & \rho_{r,v}\eta\sqrt{v_t} \\ * & * & \eta^2 \end{bmatrix} \tag{6}$$

If we set $\rho_{x,r} = \rho_{r,v} = 0$, then the model will be affine; however, as was explained in Section 2.3, we need $\rho_{x,r}$ to be non-zero. To explore the approximations made by Grzelak and Oosterlee (2011) we will first set $\rho_{r,v} = 0$, thus the only non-affine term in the covariance matrix is $\mathbf{\Sigma}_{(\mathbf{1},\mathbf{3})}$. We only need to approximate the non-affine term in $\mathbf{\Sigma}_{(\mathbf{1},\mathbf{3})}$; i.e. $\sqrt{v_t}$.

There are two approximations for $\sqrt{v_t}$ explored in Grzelak and Oosterlee (2011): a deterministic and a stochastic approach. First we explain these approaches, and then provide the extension of the deterministic approximation that allows a non-zero $\rho_{r,v}$.

3.1 Deterministic Approach

The first approach involves replacing $\sqrt{v_t}$ by its expectation, $\mathbf{E}[\sqrt{v_t}]$. As v_t is a Cox-Ingersoll-Ross (CIR) process, the expectation and variance of its square-root are given in Lemma 1 (Cox *et al.*, 1985). The resultant model will be called the HHW1 model.

Lemma 1 (Expectation and variance of the square-root of a CIR process)**.** *For any CIR process v_t, at any time $t > 0$, the expectation and variance of $\sqrt{v_t}$ are:*

$$\mathbf{E}[\sqrt{v_t}] = \sqrt{2c(t)}e^{-\lambda(t)/2}\sum_{k=0}^{\infty}\frac{\lambda(t)^k}{k!2^k}\frac{\Gamma(\frac{1+d}{2}+k)}{\Gamma(\frac{d}{2}+k)},$$

and

$$\mathbf{V}ar[\sqrt{v_t}] = c(t)(d+\lambda(t)) - \left(\mathbf{E}[\sqrt{v_t}]\right)^2,$$

where

$$c(t) = \frac{1}{4\kappa}\gamma^2(1-e^{-\kappa t}), \quad d = \frac{4\kappa\bar{v}}{\gamma^2}, \quad \lambda(t) = \frac{4\kappa v_0 e^{-\kappa t}}{\gamma^2(1-e^{-\kappa t})},$$

and $\Gamma(x)$ is the gamma function.

Proof. See Grzelak and Oosterlee (2011:Lemma 3.1) for the proof. □

However, this expectation involves an infinite sum and is rather complicated and expensive to calculate.[5] Therefore, Grzelak and Oosterlee (2011) propose that $\mathbf{E}[\sqrt{v_t}]$ is estimated by the *delta method*, which is a statistical method for approximating the expectation of a function of a random variable. The expectation of the function is approximated by a first order Taylor expansion of the function, centered around the mean of the random variable (Oehlert, 1992).[6] Applying the *delta method* to approximate $\mathbf{E}[\sqrt{v_t}]$ yields:

$$\mathbf{E}[\sqrt{v_t}] \approx \sqrt{c(t)(\lambda(t)-1) + c(t)d + \frac{c(t)d}{2(d+\lambda(t))}} =: \Lambda_t, \tag{7}$$

[5]Although the value of $\lambda(t)$ depends on the values chosen for the parameters, it was found that during calibration, $\lambda(t)$ was more often than not greater than 2.

[6]See Oehlert (1992) for more information on the *delta method*.

where $c(t)$, d and $\lambda(t)$ are as defined in Lemma 1.[7]

It is noted that the calculation of Λ_t is still non-trivial, and hence a further approximation is given as:

$$\widetilde{\Lambda}_t := a + be^{-ct} \approx \mathbf{E}[\sqrt{v_t}], \tag{8}$$

where a, b and c are constants (Grzelak and Oosterlee, 2011). Approximating Λ_t with the equation given in (8) performs very well in numerical tests conducted by Grzelak and Oosterlee (2011:p.8).

The values for these constants should be determined such that the difference between $\widetilde{\Lambda}_t$ and Λ_t is minimised. Grzelak and Oosterlee (2011) propose finding these values by matching the functions in their limits: i.e. for $t \to \infty$, $t \to 0$ and $t \to 1$. From an analysis of Λ_t and $\widetilde{\Lambda}_t$ in these limits, we have the following values for a, b and c:

$$\begin{aligned} a &= \sqrt{\bar{v} - \frac{\gamma^2}{8\kappa}}, \\ b &= \sqrt{v_0} - a, \\ c &= -\ln\left(\frac{\Lambda_1 - a}{b}\right). \end{aligned} \tag{9}$$

Note that if $\bar{v} < \frac{\gamma^2}{8\kappa}$, then a is undefined. However, if the Feller condition holds, then there will not be a problem.[8] If the Feller condition is not satisfied, we cannot use the approximation given by (8), but can still use the approximation $\sqrt{v_t} \approx \mathbf{E}[\sqrt{v_t}] \approx \Lambda_t$, where Λ_t is given in (7). This is important to note, as the Feller condition is often not satisfied in practice.

In summary, we have solved the affinity issue by replacing $\sqrt{v_t}$ with its expectation, which is approximated by Λ_t. This function is then replaced by another function, $\widetilde{\Lambda}_t$, which approximates Λ_t in limits of t. This results in $\sqrt{v_t} \approx \widetilde{\Lambda}_t$, where $\widetilde{\Lambda}_t$ is as defined in (8) and a, b and c are given by (9).

[7]The application of the *delta method* to approximate $\mathbf{E}[\sqrt{v_t}]$ can be found in Grzelak and Oosterlee (2011:Lemma 3.2.).

[8]The Feller condition, $2\kappa\bar{v} > \gamma^2$, ensures that the volatility process v_t of the Heston model will be strictly greater than zero (Feller, 1951).

3.1.1 Characteristic Function

In order to further simplify the problem such that an analytical form of the ChF can be found, Grzelak and Oosterlee (2011) set θ_t to a constant θ. We will follow this approach as our main goal is implementing and testing the model rather than calibrating it.[9] This means that we will not be able to reprice the current yield curve with our model. However, as we are only pricing relatively simple European options to test this model, this is not an issue. A time varying θ is needed when pricing a series of cash flows or other complex contingent claims, and, it is noted by Wang (2011) and Grzelak and Oosterlee (2011) that numerical integration will need to be used with a time varying θ, resulting in a slower model.

The ChF for the HHW1 model, with $\rho_{r,v} = 0$ and with the non-affine term $\eta\rho_{x,r}\sqrt{v_t}$ approximated as $\eta\rho_{x,r}\mathbf{E}[\sqrt{v_t}]$ is:

$$\phi_{HHW1}(u,\tau) = \exp[\mathcal{A}(u,\tau) + \mathcal{B}(u,\tau)x_t + \mathcal{C}(u,\tau)r_t + \mathcal{D}(u,\tau)v_t], \qquad (10)$$

where, $\tau = T - t$ and with $\mathcal{A}(u,0) = \mathcal{C}(u,0) = \mathcal{D}(u,0) = 0$ and $\mathcal{B}(u,0) = ui$ as the respective boundary conditions (Grzelak and Oosterlee, 2011). Lemma 2 gives the analytical solution for the functions $\mathcal{A}(u,\tau)$, $\mathcal{B}(u,\tau)$, $\mathcal{C}(u,\tau)$ and $\mathcal{D}(u,\tau)$.

Lemma 2 (Solution to the characteristic function of HHW1 model)**.** *The solution to the functions $\mathcal{A}(u,\tau)$, $\mathcal{B}(u,\tau)$, $\mathcal{C}(u,\tau)$ and $\mathcal{D}(u,\tau)$ in $\phi_{HHW1}(u,\tau)$ are given by:*

$$\begin{aligned}
\mathcal{B}(u,\tau) &= iu, \\
\mathcal{C}(u,\tau) &= \frac{iu-1}{\lambda}(1-e^{-\lambda\tau}), \\
\mathcal{D}(u,\tau) &= \frac{1-e^{-D_1\tau}}{\gamma^2(1-ge^{-D_1\tau})}(\kappa - \gamma\rho_{x,v}iu - D_1), \\
\mathcal{A}(u,\tau) &= \lambda\theta\mathcal{I}_1(\tau) + \kappa\bar{v}\mathcal{I}_2(\tau) + \frac{1}{2}\eta^2\mathcal{I}_3(\tau) + \eta\rho_{x,r}\mathcal{I}_4(\tau),
\end{aligned}$$

[9] This is also the approach taken by other authors investigating the HHW model.

where, $D_1 = \sqrt{(\gamma\rho_{x,v}iu - \kappa)^2 - \gamma^2 iu(iu-1)}$, $g = \frac{\kappa - \gamma\rho_{x,v}iu - D_1}{\kappa - \gamma\rho_{x,v}iu + D_1}$, *the* $\mathcal{I}_i(\tau)$*s are integrals with solutions given below, and the other parameters are as defined in system (4).*

The integrals $\mathcal{I}_i(\tau)$ $(i = 1, 2, 3)$ *have analytical solutions, and* $\mathcal{I}_4(\tau)$ *has a closed form solution if we use the approximation for* $\mathbf{E}[\sqrt{v_t}] \approx \widetilde{\Lambda}_t$*:*

$$
\begin{aligned}
\mathcal{I}_1(\tau) &= \frac{1}{\lambda}(iu-1)\left(\tau + \frac{1}{\lambda}(e^{-\lambda\tau} - 1)\right), \\
\mathcal{I}_2(\tau) &= \frac{\tau}{\gamma^2}(\kappa - \gamma\rho_{x,v}iu - D_1) - \frac{2}{\gamma^2}\ln\left(\frac{1 - ge^{-D_1\tau}}{1-g}\right), \\
\mathcal{I}_3(\tau) &= \frac{1}{2\lambda^3}(i+u)^2(3 + e^{-2\lambda\tau} - 4e^{-\lambda\tau} - 2\lambda\tau), \\
\mathcal{I}_4(\tau) &= -\frac{1}{\lambda}(iu + u^2)\int_0^\tau \mathbf{E}[\sqrt{v_{T-s}}](1 - e^{-\lambda s})ds \\
&= -\frac{1}{\lambda}(iu + u^2)\left[\frac{b}{c}(e^{-ct} - e^{-cT}) + a\tau + \frac{a}{\lambda}(e^{-\lambda\tau} - 1) + \frac{b}{c-\lambda}e^{-cT}\left(1 - e^{-\tau(\lambda - c)}\right)\right].
\end{aligned}
$$

Proof. The proof can be found in Grzelak and Oosterlee (2011:Lemma 3.6). □

3.2 Stochastic Approach

The second estimation approach is to replace $\sqrt{v_t}$ with a stochastic process equal to it in distribution. This model will be called the HHW2 model.

$\sqrt{v_t}$ is approximately normally distributed (Grzelak and Oosterlee, 2011). This can be proven by using a centralised chi-squared distribution to approximate v_t, as shown in Grzelak and Oosterlee (2011). The process $\sqrt{v_t}$ is not twice differentiable at the origin and we therefore cannot find its dynamics by application of Ito's Lemma (Jackel, 2004).

The process ξ_t, which is equivalent in distribution to $\sqrt{v_t}$, is introduced. ξ_t is an Ito process and follows the following dynamics:

$$
d\xi_t = \quad \mu_t dt + \quad \sigma_t dW_t^v, \quad \xi_0 = \sqrt{v_0}, \tag{11}
$$

with μ_t and σ_t given by:[10]

$$\begin{aligned}
\mu_t &= \frac{d}{dt}\mathbf{E}[\sqrt{v_t}] \\
&= \frac{\Gamma(\frac{1+d}{2})}{2\sqrt{2c(t)}}\left[{}_1\dot{F}_1\left(-\frac{1}{2}, \frac{d}{2}, -\frac{\lambda(t)}{2}\right)\frac{1}{2}\gamma^2 e^{-\kappa t}\right. \\
&\quad \left. + {}_1\dot{F}_1\left(\frac{1}{2}, \frac{2+d}{2}, -\frac{\lambda(t)}{2}\right)\frac{v_0\kappa}{1-e^{\kappa t}}\right], \qquad (12) \\
\sigma_t &= \sqrt{\frac{d}{dt}\mathbf{V}ar[\sqrt{v_t}]} \\
&= \sqrt{\kappa(\bar{v}-v_0)e^{-\kappa t} - 2\mu_t\mathbf{E}[\sqrt{v_t}]}.
\end{aligned}$$

In (12), $c(t)$, d, $\lambda(t)$ and $\mathbf{E}[\sqrt{v_t}]$ are as defined in Lemma 1, and ${}_1\dot{F}_1(a,b,z) := \frac{{}_1F_1(a,b,z)}{\Gamma(b)}$, where ${}_1F_1(a,b,z)$ is a confluent hypergeometric function. These are exact expressions, however; they are not simple or cheap to compute (Grzelak and Oosterlee, 2011).

Adding this process to the HHW system of equations (4) results in:

$$\begin{aligned}
dx_t &= \left(r_t - \frac{1}{2}v_t\right)dt + \sqrt{v_t}dW_t^x, \quad x_t = \ln(S_t), S_0 > 0, \\
dv_t &= \kappa(\bar{v}-v_t)dt + \gamma\sqrt{v_t}dW_t^v, \quad v_0 > 0, \\
dr_t &= \lambda(\theta_t - r_t)dt + \eta dW_t^r, \quad r_0 > 0, \\
d\xi_t &= \mu_t dt + \sigma_t dW_t^v, \quad \xi_0 = \sqrt{v_0},
\end{aligned} \qquad (13)$$

where all parameters are as defined previously.

3.2.1 Characteristic Function

The ChF for the HHW2 model, with $\rho_{r,v} = 0$ and the non-affine term approximated by the stochastic process ξ_t, is:

$$\phi_{HHW2}(u,\tau) = \exp[\mathcal{A}(u,\tau) + \mathcal{B}(u,\tau)x_t + \mathcal{C}(u,\tau)r_t + \mathcal{D}(u,\tau)v_t + \mathcal{E}(u,\tau)\xi_t], \qquad (14)$$

where $\tau = T - t$ and with $\phi_{HHW2}(u,0) = ui$ as the boundary condition (Grzelak and Oosterlee, 2011). As with the HHW1 model, we will use a

[10] See Grzelak and Oosterlee (2011) for the derivation.

constant θ. The functions $\mathcal{B}(u,\tau)$, $\mathcal{C}(u,\tau)$ and $\mathcal{D}(u,\tau)$ are given in Lemma 2. $\mathcal{A}(u,\tau)$ and $\mathcal{E}(u,\tau)$ are complicated functions whose ODEs are given in Lemma 3 and can be solved numerically.

Solving these ODEs numerically for every u involves evaluating them many times, which is computationally inefficient: we therefore exclude this model from further investigation.[11] Faster methods of solving the ODEs or approximating them with analytical solutions need to be found before it is feasible to continue an investigation into this model.

Lemma 3 (ODEs of the HHW2 model)**.** *For all $u \in \Re$ and $\tau = T - t > 0$, the functions $\mathcal{A}(u,\tau)$ and $\mathcal{E}(u,\tau)$ satisfy the below equations. (Note: we have dropped the u in $\mathcal{A}(u,\tau)$ for notational brevity; i.e. we write $\mathcal{A}(u,\tau) =: \mathcal{A}_\tau$)*

$$\mathcal{E}'_\tau = \rho_{x,r}\eta\mathcal{B}_\tau\mathcal{C}_\tau + \sigma_t\rho_{x,v}\mathcal{B}_\tau\mathcal{E}_\tau + \gamma\sigma_t\mathcal{D}_\tau\mathcal{E}_\tau, \qquad \mathcal{E}_0 = 0$$
$$\mathcal{A}'_\tau = \kappa\bar{v}\mathcal{D}_\tau + \lambda\theta\mathcal{C}_\tau + \mu_t\mathcal{E}_\tau + \frac{\eta^2\mathcal{C}_\tau^2}{2} + \frac{\sigma_t^2\mathcal{E}_\tau^2}{2}, \quad \mathcal{A}_0 = 0$$

where all other variables are as defined in the HHW2 model (see system (13)).

Proof. The proof can be found in Grzelak and Oosterlee (2011:Lemma 4.2). □

3.3 Full correlation matrix

This is an extension of the HHW1 model to include a non-zero correlation between the interest rate and the volatility process. The resulting model is called the HHW3 model, and has the full instantaneous covariance matrix given by (6).

There is an affinity issue in two elements of this matrix: namely $\mathbf{\Sigma}_{(\mathbf{1},\mathbf{3})} = \rho_{x,r}\eta\sqrt{v_t}$ and $\mathbf{\Sigma}_{(\mathbf{2},\mathbf{3})} = \rho_{r,v}\eta\sqrt{v_t}$, both of which have $\sqrt{v_t}$ as the root of this issue. In both cases we can approximate $\sqrt{v_t}$ with $\mathbf{E}[\sqrt{v_t}]$, using the same

[11] A similar approach has been taken by many papers exploring this model; for example Wang (2011).

approach as in the HHW1 model (Grzelak and Oosterlee, 2011). This solves the affinity issue.

3.3.1 Characteristic Function

The ChF for the HHW3 model is:

$$\phi_{HHW3}(u,\tau) = \exp[\hat{\mathcal{A}}(u,\tau) + \hat{\mathcal{B}}(u,\tau)x_t + \hat{\mathcal{C}}(u,\tau)r_t + \hat{\mathcal{D}}(u,\tau)v_t], \qquad (15)$$

where $\tau = T - t$ and with $\phi_{HHW3}(u,0) = ui$ as the boundary condition (Grzelak and Oosterlee, 2011). Lemma 4 gives the analytical solution for the functions $\hat{\mathcal{A}}(u,\tau)$, $\hat{\mathcal{B}}(u,\tau)$, $\hat{\mathcal{C}}(u,\tau)$ and $\hat{\mathcal{D}}(u,\tau)$.

Lemma 4 (Solution to the characteristic function of HHW3 model)**.** *The solutions to the functions $\hat{\mathcal{A}}(u,\tau)$, $\hat{\mathcal{B}}(u,\tau)$, $\hat{\mathcal{C}}(u,\tau)$ and $\hat{\mathcal{D}}(u,\tau)$ in $\phi_{HHW3}(u,\tau)$ are given by:*

$$\begin{aligned}
\hat{\mathcal{B}}(u,\tau) &= \mathcal{B}(u,\tau),\\
\hat{\mathcal{C}}(u,\tau) &= \mathcal{C}(u,\tau),\\
\hat{\mathcal{D}}(u,\tau) &= \mathcal{D}(u,\tau),\\
\hat{\mathcal{A}}(u,\tau) &= \mathcal{A}(u,\tau) + \rho_{r,v}\gamma\eta\int_0^\tau \mathbf{E}[\sqrt{v_{T-s}}]\hat{\mathcal{C}}(u,s)\hat{\mathcal{D}}(u,s)ds,
\end{aligned}$$

where, $\mathcal{A}(u,\tau)$, $\mathcal{B}(u,\tau)$, $\mathcal{C}(u,\tau)$ and $\mathcal{D}(u,\tau)$ are as defined in Lemma 2 and the other parameters are as defined in system (4).

Proof. The proof can be found in Grzelak and Oosterlee (2011:Lemma B.1). □

4 Implementation in MATLAB

This section discusses the implementation of the models in MATLAB, and presents some pricing results.

4.1 Fourier-Cosine expansion

We will use Fourier-Cosine Expansion (COS) pricing as introduced by Fang and Oosterlee (2008). COS pricing has the advantage that options of many different strikes can be priced simultaneously. It can also calculate vectors of deltas and gammas at the same time as this vector of prices. The COS method is now the industry standard technique for Fourier pricing (McWalter, 2013).

Theorem 1 is used to price options with the COS method.

Theorem 1 (Pricing under the Fourier-Cosine method)**.** *The value of an option $v(x,t)$ at time t, and with $x = \ln(S_t/K)$, is given by:*

$$v(x,t) \approx e^{-r(T-t)} \sum_{k=0}^{N-1}{}' \Re\left\{ \phi\left(\frac{k\pi}{b-a}; x\right) e^{-ik\pi\frac{a}{b-a}} \right\} V_k. \tag{16}$$

Where $\sum'$ denotes the first term in the summation is multiplied by $1/2$; $\Re\{y\}$ is the Real part of y; N represents the number of cosine terms that are summed; $\phi(u;x)$ is the the ChF of x; a and b are the lower and upper integration bounds; V_k are the cosine coefficients for the option payoff; S_t is the current stock price; K is the option strike price; r is the risk free rate and T is the option maturity.

In some cases, such as with the Heston and HHW models, the ChF can be re-parameterised, for simplification, as $\phi(u;x) = \varphi(u)e^{iux}$ and then (16) becomes:

$$v(x,t) \approx e^{-r(T-t)} \sum_{k=0}^{N-1}{}' \Re\left\{ \varphi\left(\frac{k\pi}{b-a}\right) e^{ik\pi\frac{x-a}{b-a}} \right\} V_k. \tag{17}$$

Proof. The following is proof is adapted from Fang and Oosterlee (2008) by McWalter (2013). For more details, see the original paper by Fang and Oosterlee (2008).

The Fourier-Cosine series expansion for a function $f(\theta)$ with support on $[0, \pi]$ is:

$$f(\theta) = \sum_{k=0}^{\infty}{}' A_k \cos(k\theta),$$

where $\sum'$ denotes the first term in the summation is multiplied by $1/2$ and the cosine coefficients are:

$$A_k := \frac{2}{\pi} \int_0^{\pi} f(\theta) \cos(k\theta) d\theta.$$

Using the change of variables:

$$\theta := \frac{x-a}{b-a}\pi \qquad \Rightarrow \qquad x := \frac{b-a}{\pi}\theta + a,$$

we obtain the cosine expansion for a function supported by a continuous interval $[a, b] \in \mathbb{R}$. Thus we have:

$$f(x) = \sum_{k=0}^{\infty}{}' A_k \cos\left(k\pi \frac{x-a}{b-a}\right),$$

with:

$$A_k := \frac{2}{b-a} \int_a^b f(x) \cos\left(k\pi \frac{x-a}{b-a}\right) dx.$$

The value at time t of a European option using risk-neutral pricing is:

$$v(x,t) = e^{-r(T-t)} \mathbf{E}^Q[v(y,t)|x] = e^{-r(T-t)} \int_{-\infty}^{\infty} v(y,T) f(y|x) dy,$$

where $x = \ln(S_t/K)$, $y = \ln(S_T/K)$, $v(y,T)$ is the option payoff and $f(y|x)$ is the appropriate risk-neutral density function. We can truncate this integral to an acceptable range $[a, b]$ because the density quickly decreases to zero as $y \to \pm\infty$, resulting in the approximation:

$$v(x,t) \approx e^{-r(T-t)} \int_a^b v(y,T) f(y|x) dy.$$

Then, because it is difficult or impossible to find an analytical expression for the density function $f(y|x)$, we replace it with its cosine expansion and simplify. We have:

$$f(y|x) = \sum_{k=0}^{\infty}{}' A_k \cos\left(k\pi \frac{y-a}{b-a}\right),$$

with:

$$\begin{aligned}
A_k &= \frac{2}{b-a}\int_a^b f(y|x)\cos\left(k\pi\frac{y-a}{b-a}\right)dy \\
&= \frac{2}{b-a}\int_a^b f(y|x)\Re\left\{e^{ik\pi\frac{y-a}{b-a}}\right\}dy \qquad \text{by Euler's formula} \\
&= \frac{2}{b-a}\Re\left\{\left(\int_a^b f(y|x)e^{ik\pi\frac{y}{b-a}}ds\right)e^{ik\pi\frac{-a}{b-a}}\right\} \\
&\approx \frac{2}{b-a}\Re\left\{\phi\left(\frac{k\pi}{b-a};x\right)e^{ik\pi\frac{-a}{b-a}}\right\}.
\end{aligned}$$

The last line follows because we approximate the integral $\int_a^b$ with $\int_{-\infty}^{\infty}$ and then notice that we have:

$$\int_{-\infty}^{\infty} f(y|x)e^{ik\pi\frac{y}{b-a}}ds,$$

which is the ChF for $f(y|x)$ by definition.

The option price can now be written as:

$$\begin{aligned}
v(x,t) &\approx e^{-r(T-t)}\int_a^b v(y,T)f(y|x)dy \\
&\approx e^{-r(T-t)}\int_a^b v(y,T)\sum_{k=0}^{\infty}{}'\frac{2}{b-a}\Re\left\{\phi\left(\frac{k\pi}{b-a};x\right)e^{ik\pi\frac{-a}{b-a}}\right\}\cos\left(k\pi\frac{y-a}{b-a}\right)dy.
\end{aligned}$$

Swapping the order of integration and summation and defining:

$$V_k := \frac{2}{b-a}\int_a^b v(y,T)\cos\left(k\pi\frac{y-a}{b-a}\right)dy,$$

we have:

$$v(x,t) \approx e^{-r(T-t)}\sum_{k=0}^{\infty}{}'\Re\left\{\phi\left(\frac{k\pi}{b-a};x\right)e^{ik\pi\frac{-a}{b-a}}\right\}V_k.$$

Finally, if we sum only the first $N \in \mathbb{N}$ terms we have the result (16). $\square$

In (16) and (17), the evaluation of the ChF and of the cosine coefficients are separate: therefore, K (hence x) could be a vector of strike prices. This is the source of the pricing advantages for this method.

The V_k, $k \in \mathbb{N}$, are the cosine coefficients of the option payoff. There are analytical solutions for the V_k for several option contracts. The cosine coefficients for European call and put options are given in Lemma 5, but first we need two basic results given by Proposition 1.

Proposition 1 (Cosine coefficients for e^x and 1)**.** *The cosine coefficients for e^x on $[c,d] \subset [a,b]$ are:*

$$\begin{aligned}\chi_k(c,d) &= \int_c^d e^x \cos\left(k\pi\frac{x-a}{b-a}\right) dx \\ &= \frac{1}{1+\left(\frac{k\pi}{b-a}\right)^2}\Bigg\{ \cos\left(k\pi\frac{d-a}{b-a}\right)e^d - \cos\left(k\pi\frac{c-a}{b-a}\right)e^c \\ &\quad + \frac{k\pi}{b-a}\left[\sin\left(k\pi\frac{d-a}{b-a}\right)e^d - \sin\left(k\pi\frac{c-a}{b-a}\right)e^c\right]\Bigg\}.\end{aligned}$$

The cosine coefficients for 1 on $[c,d] \subset [a,b]$ are:

$$\begin{aligned}\psi_k(c,d) &= \int_c^d 1\cos\left(k\pi\frac{x-a}{b-a}\right) dx \\ &= \begin{cases} \frac{b-a}{k\pi}\left[\sin\left(k\pi\frac{d-a}{b-a}\right) - \sin\left(k\pi\frac{c-a}{b-a}\right)\right] & n > 0, \\ d-c & n = 0. \end{cases}\end{aligned}$$

Proof. By basic integration rules. □

Lemma 5 (Cosine-coefficients, V_k, for vanilla European options)**.** *For a call option we have (for $k \in \mathbb{N}$):*

$$V_k^{call} = \frac{2}{b-a}K(\chi_k(0,b) - \psi_k(0,b)),$$

where K is the option strike price, and $[a,b]$ the integration bounds from the Fourier-Cosine expansion, with $\chi_k(c,d)$ and $\psi_k(c,d)$ as defined in Proposition 1.

Similarly, for a put option we have (for $k \in \mathbb{N}$):

$$V_k^{put} = \frac{2}{b-a} K(-\chi_k(a,0) + \psi_k(a,0)).$$

Proof. The payoff of a call option is $v(y,T) = \max(S_T - K, 0)$. Using the transformation $x = \ln(S_t/K)$ and $y = \ln(S_T/K)$, we have:

$$\begin{aligned} v(y,T) &= \max(S_T - K, 0) \\ &= \max(K(e^y - 1), 0) \\ &= K(e^y - 1)\mathbb{I}_{y>0}. \end{aligned}$$

Then, using the definition of V_k, and simplifying we get:

$$\begin{aligned} V_k &= \frac{2}{b-a} \int_a^b v(y,T) \cos\left(k\pi \frac{y-a}{b-a}\right) dy \\ &= \frac{2}{b-a} \int_a^b K(e^y - 1)\mathbb{I}_{y>0} \cos\left(k\pi \frac{y-a}{b-a}\right) dy \\ &= \frac{2}{b-a} \int_0^b K(e^y - 1) \cos\left(k\pi \frac{y-a}{b-a}\right) dy. \end{aligned}$$

Using Proposition 1 we have:

$$V_k = \frac{2}{b-a} K(\chi_k(0,b) - \psi_k(0,b)),$$

as required.

The proof for the put option is similar. □

4.1.1 Integration Bounds

The COS method has exponential convergence in N, provided the upper and lower bounds of integration (i.e. the interval $[a,b]$) are chosen to be sufficiently wide (Fang and Oosterlee, 2008). Fang and Oosterlee (2008) suggest that the bounds be chosen as:

$$[a,b] := \left[c_1 - L\sqrt{c_2 + \sqrt{c_4}},\ \ c_1 + L\sqrt{c_2 + \sqrt{c_4}}\right], \tag{18}$$

where c_i is the ith cummulant of the distribution for x and L is a constant, suggested as 10.

However, the calculation of the 4th cummulant for the Heston model is rather involved and so Fang and Oosterlee (2008) propose:

$$[a,b] := \left[c_1 - 12\sqrt{|c_2|},\ \ c_1 + 12\sqrt{|c_2|}\right],$$

for the Heston bounds. Grzelak *et al.* (2012) note that the calculations of the HHW cummulants are not available, and suggest the following approximation:

$$[a,b] := \left[0 - 8\sqrt{\tau},\ \ 0 + 8\sqrt{\tau}\right],$$

where τ is the time to maturity. We will make use of this approximation.

4.1.2 The Greeks

To get the delta of an option, one simply takes the partial derivative of (17) with respect to the stock price to get:

$$\frac{dv(x,t)}{dS_t} \approx e^{-r(T-t)} \sum_{k=0}^{N-1}{}' \Re\left\{\varphi\left(\frac{k\pi}{b-a}\right) e^{ik\pi\frac{x-a}{b-a}} \frac{ik\pi}{b-a}\right\} \frac{1}{S_t} V_k,$$

because $\frac{dv}{dS} = \frac{dv}{dx}\frac{dx}{dS}$.

Again, the evaluation of the ChF is separate and the same calculations of the ChF from the price calculation can be used when solving the delta. The gamma also has this advantage and is given by:

$$\frac{d^2v(x,t)}{dS_t^2} \approx e^{-r(T-t)} \sum_{k=0}^{N-1}{}' \Re\left\{\varphi\left(\frac{k\pi}{b-a}\right) e^{ik\pi\frac{x-a}{b-a}} \left[-\frac{ik\pi}{b-a} + \left(\frac{ik\pi}{b-a}\right)^2\right]\right\} \frac{1}{S_t^2} V_k.$$

Thus, calculation of the price or the price, delta and gamma takes roughly the same time. This is another advantage of the COS pricing method.

Taking the partial derivative of (17) with respect to v_t, the variance process at time t, we get:

$$\frac{dv(x,t)}{dv_t} \approx e^{-r(T-t)} \sum_{k=0}^{N-1}{}' \Re\left\{\frac{d}{dv_t}\varphi\left(\frac{k\pi}{b-a}\right) e^{ik\pi\frac{x-a}{b-a}}\right\} V_k,$$

where under the HHW1 model we have:

$$\frac{d}{dv_t}\varphi_{HHW1}\left(\frac{k\pi}{b-a}\right) = \varphi_{HHW1}\left(\frac{k\pi}{b-a}\right)\mathcal{D}\left(\frac{k\pi}{b-a}, T-t\right),$$

and $\mathcal{D}(u,\tau)$ is as defined in Lemma 2.

Notice that this is not the vega: vega is $\frac{d}{d\sigma}$, whereas we have $\frac{d}{dv_t}$. To find the vega, we apply Ito's lemma and get $\frac{d}{d\sigma} = \frac{d}{dv_t}\frac{dv_t}{d\sigma}$. Therefore, the vega of the option is:[12]

$$\frac{dv(x,t)}{d\sigma} = 2\frac{dv(x,t)}{dv_t}.$$

Due to the fact the v_t appears in the ChF, vega does not have the advantage of being able to use the same evaluation of the ChF from the price calculation. Thus, calculating both the price and the vega takes roughly twice as long as calculating the price.

4.1.3 The HHW1 and Heston models with COS pricing

When implementing the HHW1 model, we cannot simply take the ChF as given in Lemma 2 and implement it. There are some key points to note and some changes to be made.

Firstly, the presented ChF is a discounted ChF and thus we do not need to apply any discounting when evaluating the price with the COS method (Grzelak *et al.*, 2012).[13] Contrary to this, the commonly published Heston ChF is not discounted and thus must be discounted when pricing with the COS method.

Secondly, we re-parameterise the ChF as in Theorem 1 by removing $\mathcal{B}(u,\tau) = e^{iux}$, and rewriting the ChF as:

$$\phi_{HHW1}(u,\tau) = \varphi_{HHW1}(u,\tau)e^{iux}.$$

[12] The variance process $v_t = \sigma_t^2$.

[13] Grzelak *et al.* (2012) show that a discounted ChF can be used instead of discounting in the COS method.

This allows us to use the advantages of the COS method by separating the strike price from the evaluation of the ChF. Again, the Heston ChF is often presented without the e^{iux} term and one must remember to include it in the evaluation of the COS formula; i.e. use (17) and not (16) with this model.

Furthermore, some papers neglect to mention that even though we are working in the complex plane, some of the square-roots are not meant to produce complex numbers. We refer specifically to the calculation of D_1 (or D) when evaluating the HHW (or Heston) ChFs. This square-root should be "such that its real part is nonnegative" (Fang and Oosterlee, 2008:p.8).[14]

In addition, it is noted by Fang and Oosterlee (2008) that when pricing call options with the COS method, the results are sensitive to the size of the integration range; i.e. the choice of L in (18). This is due to a call option payoff having unbounded value in S_t, and is not observed with put option payoffs, whose values are bounded by K. Thus, one should price a put option and then use the put-call parity to find the value of a call option.

Finally, in the HHW1 model when the Feller condition is not satisfied, the approximation of the function Λ_t by $\widetilde{\Lambda}_t$ is not defined, as the square-root in (9) is not defined. Thus, we do not use the approximation $\widetilde{\Lambda}_t$ given by (8) and instead evaluate the integral:

$$\begin{aligned}\mathcal{I}_4(\tau) &= -\frac{1}{\lambda}(iu+u^2)\int_0^\tau \mathbf{E}[\sqrt{v_{T-s}}](1-e^{-\lambda s})ds \\ &\approx -\frac{1}{\lambda}(iu+u^2)\int_0^\tau \Lambda_{T-s}(1-e^{-\lambda s})ds\end{aligned}$$

in Lemma 2, where the approximation of $\mathbf{E}[\sqrt{v_{T-s}}]$ by Λ_{T-s} is given by (7).

Evaluating the integral with MATLAB's `integral` command takes ± 0.06 seconds on the first evaluation and then 0.002 seconds for subsequent evaluations. This causes the pricing to be slightly slower than if the approximation was used, but makes the pricing more versatile as the Feller condition is often not satisfied in practice.

[14] This is very important to be aware of when working in MATLAB, because MATLAB will take the root of a negative value without issuing any warning.

4.1.4 Implementing the HHW1 model

Taking into consideration the changes in Section 4.1.3, we simply reproduce Lemma 2 in a MATLAB function that returns the ChF. We also have a function that returns $\frac{d}{dv_t}\varphi_{HHW1}(u)$ to allow the calculation of the vega.

We then implement a COS pricing method that calculates the option cosine coefficients for a vector 0 to $N-1$, as well as the ChF for the same vector. As per the COS method, we take the real part, multiply the coefficients and ChF, adjust the first term by 1/2 and sum to get the price. This is all done using vectorised calculations as vectorised code runs much faster in MATLAB (MATLAB, 2013).[15] Using the calculated ChF, we can multiply by the extra terms required for the delta and gamma, and apply the rest of the COS method to solve for the delta and gamma. To calculate the vega, we calculate a vector of $\frac{d}{dv_t}\varphi_{HHW1}(u)$ and apply the rest of the COS method.

For calibration we also need the model to price swaptions: thankfully these are not dependent on equity processes or parameters and thus only rely on the Hull & White (H&W) part of the model. For a H&W model, swaptions prices can be calculated analytically if one uses the Jamshidian (1989) trick. We can use this same analytical formula to calculate swaption prices under a HHW model. We implement this in MATLAB in a fairly straightforward manner, using `fzero` to solve for the strike in the Jamshidian (1989) trick.

The only thing required before we can price with this model is to decide on the number of COS terms, N, to be used for the COS pricing method: this is done in Section 4.1.6.

[15] See MATLAB (2013:Vectorization) for examples on how to vectorise code.

4.1.5 Implementing the HHW3 model

The main implementation difference between HHW1 and HHW3 is the evaluation of two integrals in the function $\hat{\mathcal{A}}$ in Lemma 4 for HHW3, whereas HHW1 only has one integral, $\mathcal{I}_4(\tau)$.

The integral in the HHW1 function is not dependent on u, which allows us to completely vectorise the implementation of the COS method in this model.

This is not possible in the HHW3 model, as the second integral:

$$\rho_{r,v}\gamma\eta\int_0^\tau \mathbf{E}[\sqrt{v_{T-s}}]\hat{\mathcal{C}}(u,s)\hat{\mathcal{D}}(u,s)ds,$$

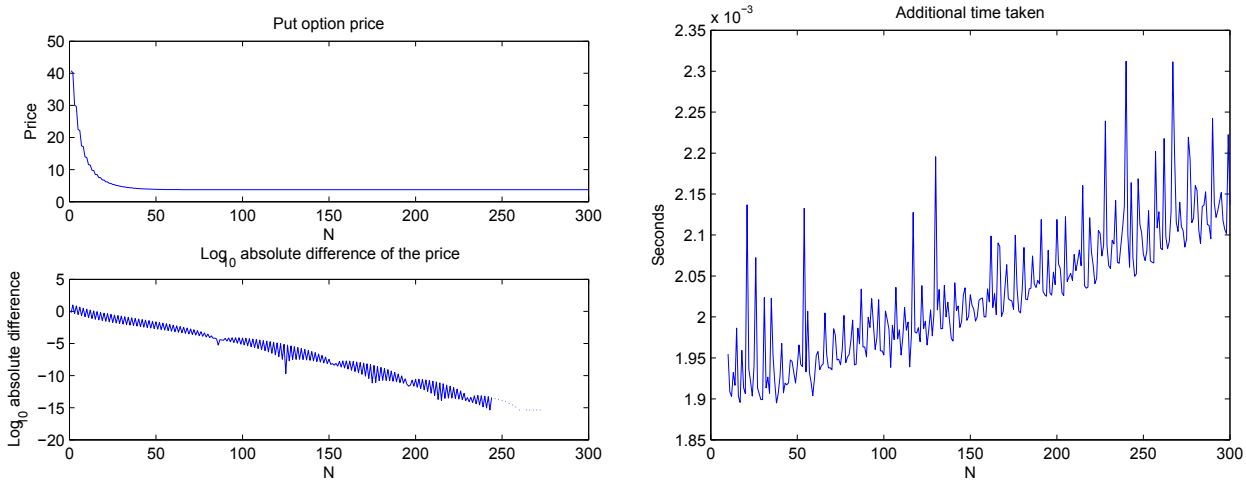
depends on u and thus a `for` loop needs to be used when calculating the COS price under this model. Due to the speed enhancements available in MATLAB for vectorised code, the HHW1 model prices options much faster (on the order of 15 - 20 times) than the HHW3 model.

4.1.6 Efficiency and Error testing

Figure 1(a) shows the price and the $\log_{10}$ absolute difference in the price of an at-the-money (ATM) European put option priced with the COS method as a function of the number of COS terms, N.[16] The price stabilises very quickly, and with $N \geq 200$, we have entered the territory of machine error as seen in the $\log_{10}$ absolute difference plot in Figure 1(a). This stabilisation takes longer with call options prices, as mentioned in Section 4.1.3. However, if we price put options and then use the put-call parity to find the call option prices, these prices will exhibit the same convergence as the put options in Figure 1(a).

There is a slight increase in the time taken to compute the prices as we increase N, but only very slight, as displayed in Figure 1(b). The actual time taken to price is not displayed in 1(b). There is a run-up time of roughly

[16]The HHW1 model was used and the model parameters are as given in Section 4.3.



(a) Price and $\log_{10}$ absolute difference

(b) Additional time taken

Figure 1: At the money put option price, $\log_{10}$ absolute difference of the price and the additional time taken to price with the COS method. Note that the time taken plot starts at $N = 10$, not $N = 1$. There is a run-up cost to pricing of 0.12 seconds, which should be added to the time taken for each N.

0.12 seconds to price one option with the COS method in MATLAB. When pricing multiple options, this run-up only happens once, and after the first 5 to 10 option prices, the next option price only takes 0.002 seconds to compute. This additional time taken is the time displayed in Figure 1(b). The reason for this is the way MATLAB automatically optimises `for` loops which were used to generate the prices of multiple options. This means that pricing a single option with $N = 1$ takes 0.1219 seconds, and with $N = 250$ takes 0.1223 seconds.

We have chosen to use $N = 200$ for all further pricing with the COS method as we can reasonably expect the price to have converged by here, and we can afford to use a high N as there is not much cost to additional cosine terms. The reason for this low additional cost is potentially due to MATLAB's vectorisation ability. This means that the same results may not be observed in another programming language.

Pricing one option takes roughly 0.122 seconds, pricing 100 takes 0.5 seconds and pricing 1000 only takes 1.06 seconds.[17] After removing the run-up time cost, the time taken increases linearly with the number of options being priced. This run-up cost is also higher when using parallel computing to do the pricing.

4.1.7 Parallelisation

The pricing problem does not admit parallelisation of the actual pricing process. In the HHW1 model, there is no `for` loop in the pricing process and therefore nothing to parallelise. As explained in Section 4.1.5, the HHW3 implementation does have a `for` loop, but it does not deal with large amounts of data. The overheads, i.e. cost of transferring the data to each core and the gathering it back after calculation, are too large to make parallelisation feasible for the HHW3 model.

However, the use of the pricing process can be parallelised. When we run hedging or calibration tests, we can run them in parallel. Calculating the prices of 10000 options (with random maturities and strikes) under the HHW1 model with the use of a normal for loop takes 20 seconds.[18]

We can make use of the very simple parallelisation offered by MATLAB with the `parfor` loop. If the problem permits it, a `for` loop can be replaced with a `parfor` loop for "instant parallelisation" (MATLAB, 2013). This allows us to price on 8 workers (on 4 cores) instead of one, and prices 10000 options in 6 seconds.

This is more than three times faster than without parallelisation. This shows some of the gains that are possible with parallelisation and the HHW models. MATLAB's `parfor` loop is one of the slower ways to introduce parallelisation, but is very simple and easy to do (MATLAB, 2013). Faster

[17] These prices are for random option parameters, and are computed in parallel.

[18] All pricing and calculations are done on an Intel i7 3630QM CPU @ 2.4GHz with 16GB ram.

options are available; including GPU processing, where one could get increases in speed of 2 to 5 times above the simple parallelisation (OPTI-NUM solutions (Pty) Ltd, 2013).

4.2 Monte Carlo

In order to compare prices, a Monte Carlo simulation of the full HHW model given by system (4) was also implemented. This was done in a fairly standard way, using a Milstein discretisation scheme to discretise the SDEs.

Due to the nature of the models, there is a non-zero probability of the variance, v_t, and interest rate, r_t, processes becoming negative when a discretisation scheme is used. There are a few methods to fix this as reviewed by Lord *et al.* (2010). We have chosen the method with the smallest discretisation bias, namely "full truncation", for the volatility process (Lord *et al.*, 2010:p.1). In this method, the process, v_t, is allowed to go negative, but when the process is used, we substitute $\max(v_t, 0)$ for v_t. For the interest rate process, we use the absolute value of r_t to create a reflecting barrier at zero.[19]

We use the variance reduction technique of antithetic sampling in all Monte Carlo simulations. This allows for a quicker rate of convergence and thus fewer simulations are required to achieve the same level of accuracy as a plain Monte Carlo scheme. For all Monte Carlo simulations, we use 50000 paths.

4.3 Results

To test the implementation of the models, call options with strike prices ranging from 50 to 150 were priced, where $S_0 = 100$. The parameters of the model are: $v_0 = 0.0175$; $r_0 = 0.07$; $\kappa = 1.5768$; $\bar{v} = 0.0398$; $\gamma = 0.0571$; $\lambda =$

[19]This was suggested by Old Mutual and is often done in practice when dealing with negative interest rate processes.

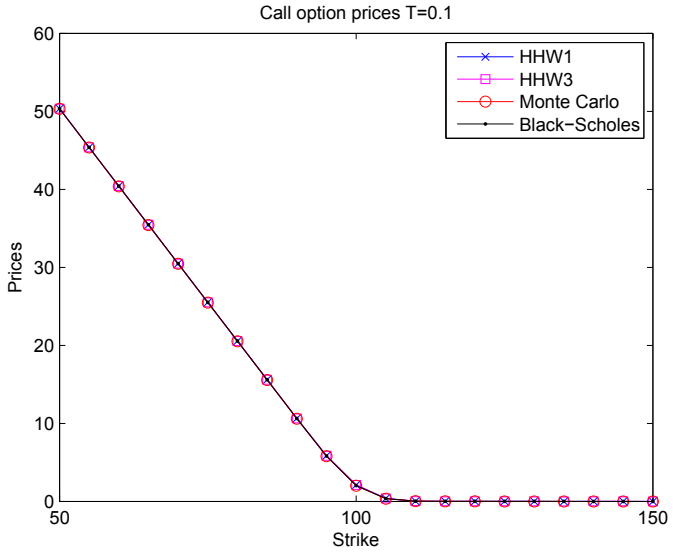


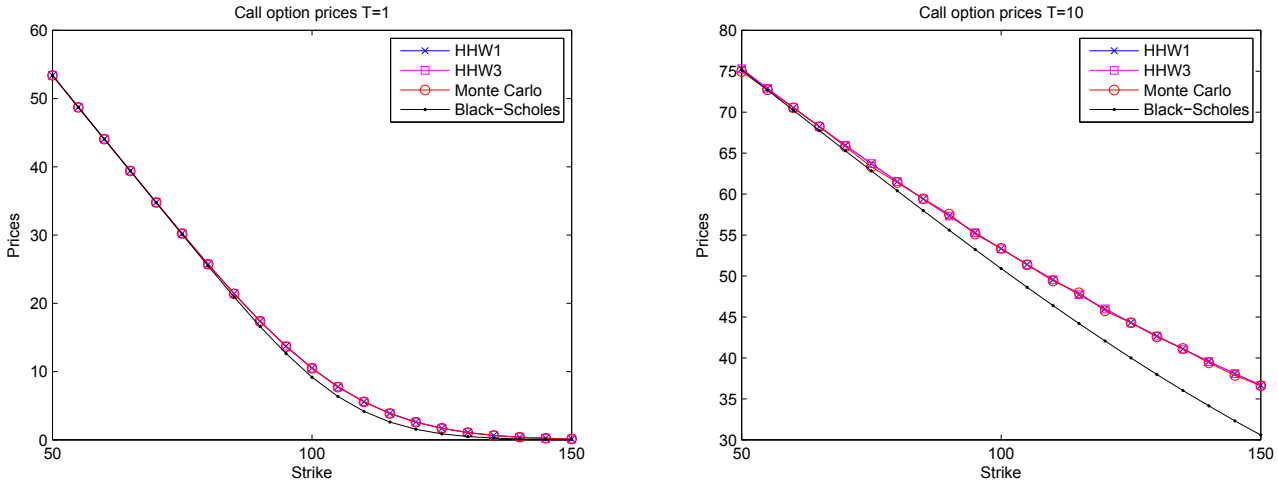
Figure 2: Call option prices for different strikes and maturity $T = 0.1$ years, under the HHW1, HHW3, Monte Carlo and Black-Scholes models.

0.05; $\theta = 0.07$; $\eta = 0.005$; $\rho_{x,v}$ = -0.5711; $\rho_{x,r} = 0.2$; $\rho_{r,v} = 0.3$, which are the same as those used by Wang (2011).

From Figure 2, it is clear that the models are pricing correctly for very short maturities. As we increase the time to maturity, the HHW prices drift away from the Black-Scholes prices, as observed in Figure 3. This is because the further away the maturity of the option, the longer stochastic effects have to act on the variables. This will result in an increase in call option prices, as these options have limited downside and unlimited upside potential. As expected, out-the-money (OTM) option prices deviate less than in-the-money (ITM) prices.[20]

The HHW1 and HHW3 models perform almost identically and the Monte Carlo prices follow quite closely. Table 5 and Table 6 in Appendix B present the numerical results of the pricing.

[20]OTM options require large movements in stock prices to increase in value, whereas ITM options will be more sensitive to smaller stock movements.



(a) T = 1 (b) T = 10

Figure 3: Call option prices for different strikes and maturities: $T = 1$ and $T = 10$ years, under the HHW1, HHW3, Monte Carlo and Black-Scholes models.

5 Calibration

When calibrating the HHW model, one should first calibrate the interest rate component and then, holding those parameters constant, calibrate the equity and stochastic volatility components of the model (Grzelak and Oosterlee, 2011). This is the most computationally efficient way to calibrate the HHW model (Wang, 2011).

There are many papers and methods exploring the calibration of the Hull & White (H&W) model and the Heston model. The best of these could be chosen and combined quite easily, as the calibrations in the HHW model are somewhat separate. It is, however, important to remember that a model with a perfect fit to the market is still a model: an approximation of the observed data. Thus, one should not waste time and resources looking for a perfect fit, but rather improve the underlying model itself.

As calibration is not the focus of this project, a simple calibration process

was employed. This process is explained in this section and some calibration results are presented.

5.1 Data

The period under review is from 2 August 2010 to 31 May 2013, which is 742 trading days. The JSE Top40 index was chosen as the equity component of the model. Price data, as well as JIBAR rates, were supplied by Old Mutual (acquired from Bloomberg).

Daily zero curves were available from 2 August 2010 to 22 November 2012. From 22 November 2012 to 31 May 2013, only monthly curves were available. Weekly volatility skews were available from 2 August 2010 to 10 September 2012. These skews were available for options on the Top40, with strikes of 70 to 130 percent of spot and maturity 0.5 years and 1 year. 5-year ATM options on the Top40, and 1-year ATM swaptions were also available. From December 2012 to May 2013, only monthly skews were available and only for options with strikes at 85, 95, 100 and 110 percent of spot. All of the above data were supplied by Old Mutual.

Where daily data were not available, the inter-month (-week) data were assumed constant for the month (week). This results in two periods in the data: the first being from 2 August 2010 to 27 August 2012 (where we have daily or weekly data) and the second from 3 September 2012 to 31 May 2013 (where we only have monthly data).

The inputs to the daily calibration process are the current short rate, r_0, the current variance, v_0, and the correlation between equity prices and interest rates, $\rho_{x,r}$. The short rate was taken as the overnight rate from the zero curve. For the current variance, we used the ATM 0.5 year equity volatility from the volatility skew.

A 120-day rolling correlation of the Top40 closing prices and the closing

12-month JIBAR interest rate was used for $\rho_{x,r}$.[21] Equity and interest rate data were needed for the period 4 January 2010 to 31 May 2013 for the calculation of correlation.

5.2 Interest rate

To calibrate the H&W part of the model, we need to find λ, θ and η. We have assumed θ to be constant, representing the long-run interest rate, thus we let θ be the 10-year zero curve rate.

We find λ and η by fitting the model to swaption data. We price 1y1y, and 1y2y swaptions under the H&W model, and then find their implied volatilities by inverting Black's formula (Black, 1976).[22] We then choose parameters λ and η to minimise the error between these model implied volatilities and the market observed volatility.

MATLAB's `lsqcurvefit` function is used to do this minimisation. The function is a convenient interface to another function, `lsqnonlin` (MATLAB, 2013). `lsqcurvefit` minimises the error between the X and Y data in a least squares manner. It uses the `trust-region-reflective` algorithm, which is MATLAB's implementation of the trust region algorithm used in minimisation problems (MATLAB, 2013). To allow for a more accurate fit, the function was run on $100 \times \sigma$ instead of just σ. Other than slightly adjusting the tolerance levels, `lsqcurvefit` is used in a standard way.

5.3 Stochastic volatility

Once the H&W part of the model is calibrated, we need to find κ, $\bar{v}$, γ, $\rho_{x,v}$ and $\rho_{r,v}$. We do this with the same approach as the interest rate calibration. We price European options under the HHW model with κ, $\bar{v}$, γ, $\rho_{x,v}$ and $\rho_{r,v}$

[21] 120 days were chosen as this is roughly half a trading year. Shorter rolling windows (30 and 60 days) resulted in very volatile correlations. A window of 90 to 140 days significantly reduced the volatility of the correlation.

[22] 1y2y is swaption shorthand for a one year option on a two year swap.

varying, and all other parameters constant. Then we find the Black-Scholes implied volatility, which we compare to the market volatility.

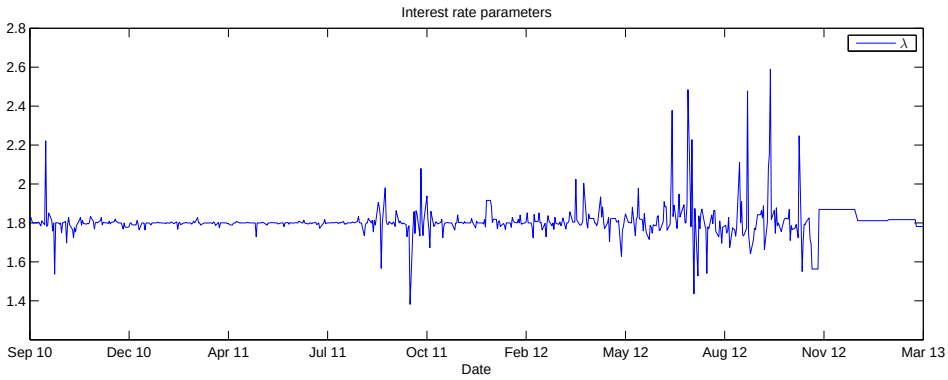
Again, we use `lsqcurvefit` to minimise the least squared error between the implied and observed volatilities. We make the same adjustment to the volatilities as in Section 5.2, and use $100 \times \sigma$.

5.4 Results

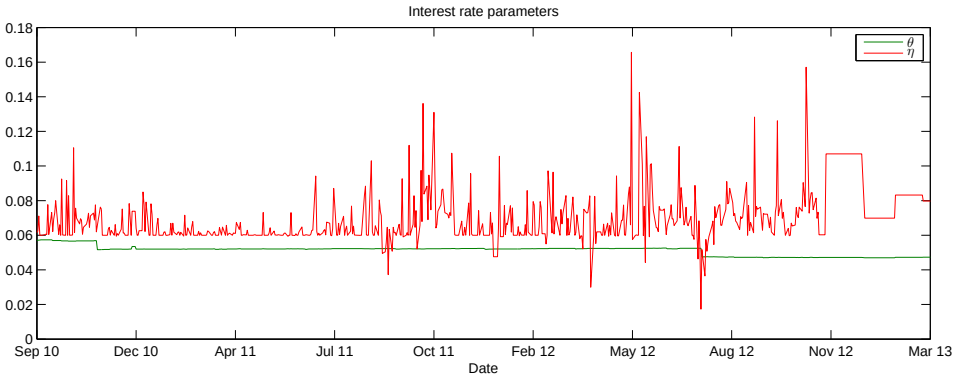
We calibrated the model to all 27 equity options and only nine in both periods with similar results. We now present the results for calibration of the HHW1 model to all 27 options in period one and to the nine (that existed) in period two. Calibrating the model to one day (with 27 equity options) takes 160 seconds on average. The fastest calibration was under 60 seconds, and 524 out of 742 calibrations were below 160 seconds. There were, however, 64 calibrations that took longer than 360 seconds: in these cases, the calibration procedure terminated without finding an acceptable minimum.

The main reason we want to implement an efficient pricing model is for calibration. During calibration, the pricing method is called 2600 times (on average). This means that for every additional second taken to price an option, the calibration process will take 44 minutes longer. In practice, calibration is at least a daily activity. If we were using a slower pricing method, such as Monte Carlo, calibration would take much too long to be a feasible daily activity. A model that takes too long to calibrate is a poor choice for hedging or risk reporting.

Another important factor when looking at calibration for hedging is how stable the parameters are over time. If the parameters are very volatile then there will be large hedge errors with this model. This is due to, for example, hedging on the first day with low parameters and then recalibrating the next day to find that the parameters are high, and that according to the model,



(a) λ



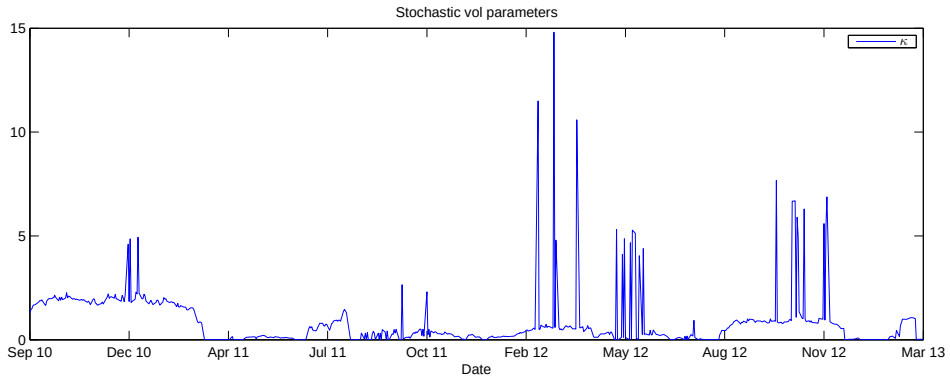
(b) θ and η

Figure 4: Calibrated daily interest rate parameters for the HHW1 model.

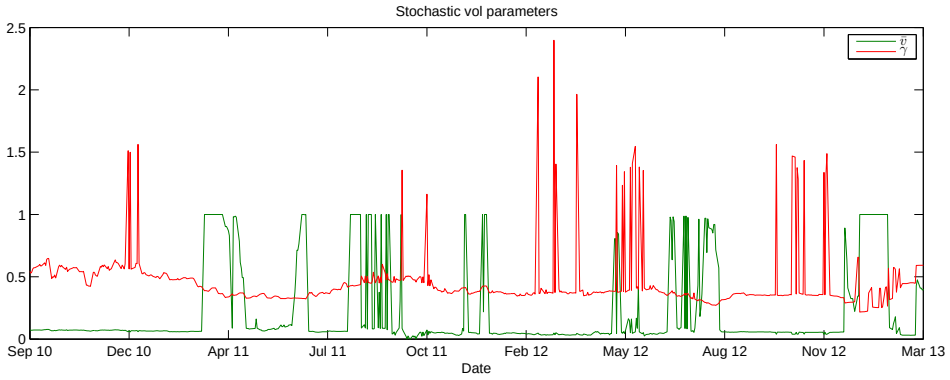
we are hedged poorly. These hedge errors are due only to fluctuations in calibration parameters and not in the market parameters. Thus, if we trade to fix them, we will be trading unnecessarily and building up trading costs.

Figure 4 shows the daily calibration results for the interest rate parameters of the model. It is clear that the interest rate parameters are quite stable, in some cases almost constant. Notice that in period two, the parameters have steps in them: this is due to the construction of the data.

The stochastic volatility parameters are more volatile than the interest rate parameters, as evident in Figure 5. There are lengths of time where even these parameters are relatively stable, thus, the model can produce stable results for these parameters. A better calibration technique may be



(a) κ



(b) $\bar{v}$ and γ

Figure 5: Calibrated daily stochastic volatility parameters for the HHW1 model.

able to increase the length of these periods of stability.

The correlation parameters $\rho_{x,r}$ and $\rho_{x,v}$ are shown in Figure 6. Under the HHW1 model, $\rho_{r,v}$ is zero and is not depicted. Comparing Figure 6 to Figure 4 and Figure 5, there is some evidence that the model parameters spike when very high or low $\rho_{x,r}$ is observed.

In Figure 7, we compare the market observed volatilities to the implied volatilities produced by the calibrated model for 2 August 2010. The calibration has worked well in the short-term, but deviates from the long-term market volatility. On other days, the calibrated curve deviated in the short-term but matched the long-term well. This shows that the model has the potential to calibrate to observed market surfaces if we use a better calibra-

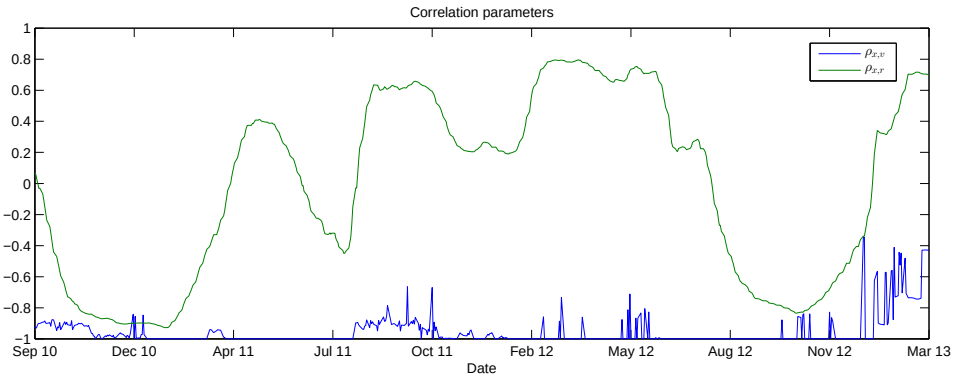
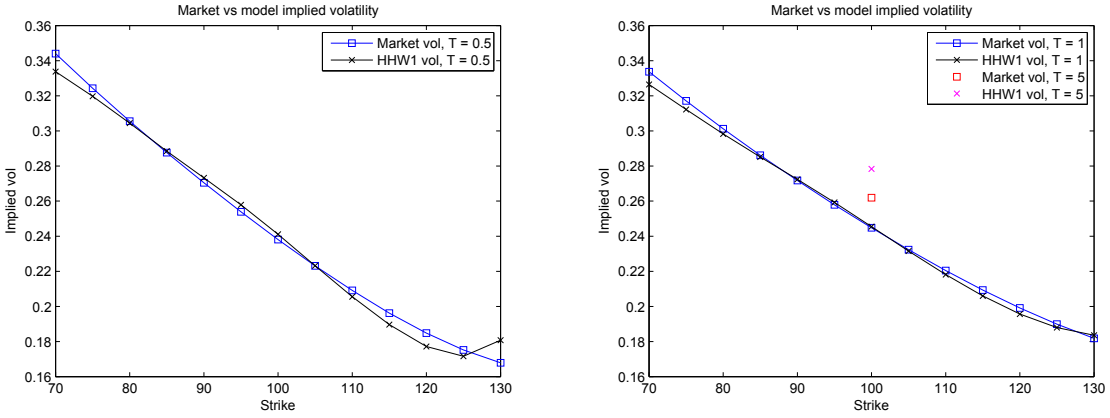


Figure 6: Calibrated daily correlation parameters for the HHW1 model.



(a) T = 0.5

(b) T = 1 and T = 5

Figure 7: Comparison of model implied volatility with market observed volatility on 2 August 2010.

tion technique.

5.5 HHW1 versus HHW3

In Section 5.4, we only present results for the HHW1 model because it is quicker to use this model, as discussed in Section 4.1.5. We have, however, calibrated the HHW3 model to 149 days in 5-day intervals to compare the HHW1 and HHW3 calibration results. Calibrating the HHW3 model took 10 times longer than the HHW1 model. Although this is much slower than

the HHW1 model, it is still fast enough to be considered a possible model choice for hedging long-dated equity options.

To speed up the HHW3 calibration process, we first calibrate a HHW1 model to the data and then use these results as the initial guess for the HHW3 calibration. This works because the fitted parameters under the HHW1 and HHW3 models are similar. The residual after the fitting process was lower for the HHW3 model, which was as expected when adding an additional parameter. It could be argued that a slight reduction in the residual is not worth the additional time taken to calibrate the model. However, if we are pricing a hybrid contingent claim that depends on the correlation between the volatility and the interest rate, then we need to include this parameter. Thus, the inclusion of non-zero correlation between the stochastic volatility process and the stochastic interest rate process may be justified.

6 Hedging Analysis

In this section, we look at hedging with the HHW1 model.[23] The hedging results for the HHW1 model, as opposed to the HHW3 model, are presented, as calibration is much quicker and we can therefore run more back-tests with this model. The same simulated tests, as described below, and comparable (but less extensive) back-tests were run under the HHW3 model and the results were similar to those of the HHW1 model. We first look specifically at delta hedging by comparing different model choices to different scenarios. Then we compare delta, delta-gamma and delta-vega hedging methods.

To simulate the world scenarios, we use the same discretisation scheme as was used in the Monte Carlo simulations for pricing.[24] 2000 realisations were simulated and used by all hedges in each world scenario. Therefore, direct comparison of the different hedging methods is possible. We calculate the prices and Greeks of the options using the COS method for the Heston and HHW1 models, and the analytical formulas for the Black-Scholes (BS) model.

Unless otherwise specified, we use the model parameters presented in Section 4.3 for the HHW1 model. The Heston model uses the same parameters where applicable. Where constant interest rates or volatility is assumed, the parameters for r_0 and v_0 (respectively) are used as the constants. Daily hedging is then simulated for a year. In these hedging simulations, the option that is being hedged is a long ATM 20-year call option. The current stock price is 100 and the price of this option at time zero, under the HHW1 model, is R76.3005. The 5-year ATM call option, that is used as a hedge instrument, has a time zero value of R33.9304.

[23] All of the hedging is done using the methods described by Kienitz and Wetterau (2012).

[24] See Section 4.2.

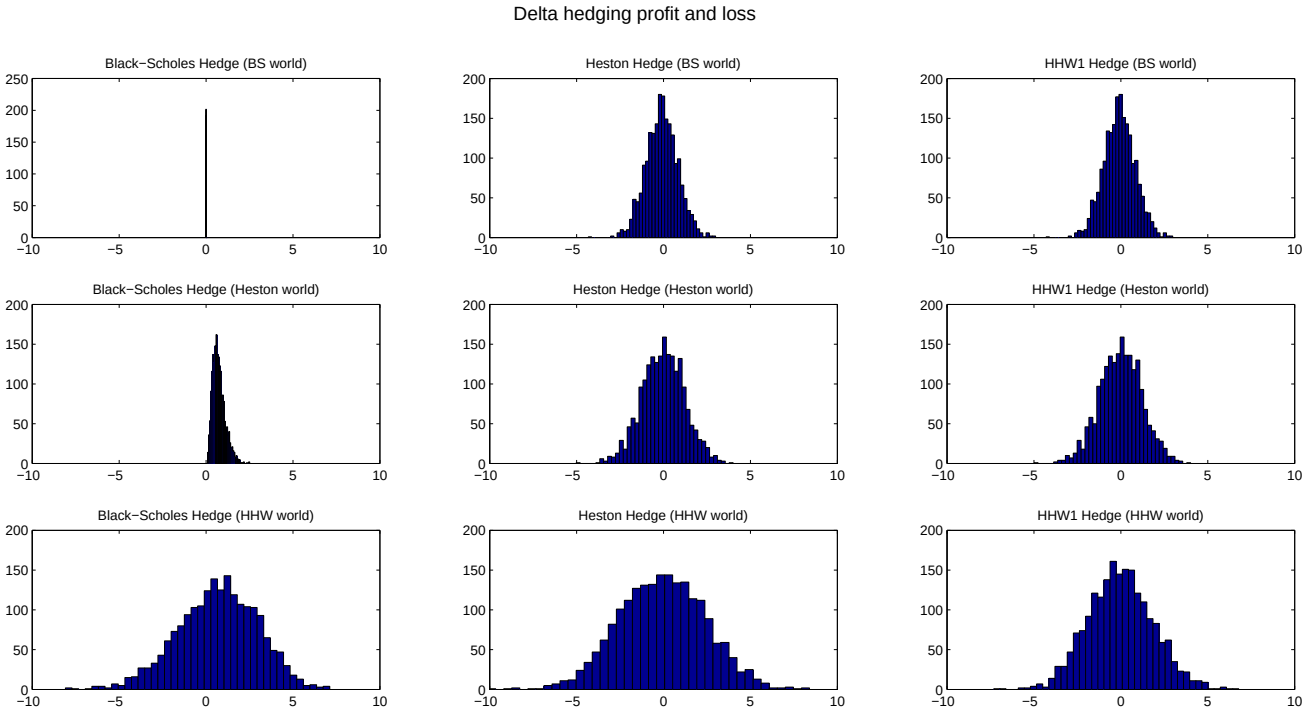


Figure 8: Daily delta hedging profit and loss (PnL) for a 20-year ATM call option. PnL is given for different hedging model choices (columns) and different world scenarios (rows).

6.1 Delta hedging

To start, we explore delta hedging with the BS, Heston and HHW1 models. There are three models under consideration and we assume there are three possible ways for the world scenarios to be generated. Thus, there are a total of nine different situations, depicted in Figure 8. A situation, for example, is that the world is generated with BS and our hedging model choice is HHW1, or that the world is Heston and we hedge with a BS model. If there is stochastic volatility in the world and we want to look at the impact of assuming constant volatility, then we are interested in the situation where the world is Heston or HHW and the model choice is BS.

In Figure 8 we show histograms of the Profit and Loss (PnL) due to hedging an ATM 20-year call option. The first row in the figure is for the

BS world, and second row in the figure is for the Heston world, where we have stochastic volatility. In both cases, simple BS hedging produces the best results: therefore, the assumption of constant volatility does not appear to result in larger losses when hedging this long-dated call option.

The third row in the figure is for a HHW world, where stochastic interest rates are added to the world. In this world, the BS and Heston hedges perform poorly; this can also be seen by the standard deviations in Table 1. The HHW1 model performs well as it has a lower standard deviation than the BS and Heston models (1.8766 versus 2.2990 and 2.4935 respectively). The HHW1 model is also centred closer to zero than the other two models, as evident by the means in Table 1. This indicates that the assumption of a constant interest rate results in hedging losses when hedging long-dated options. Thus, if there are stochastic interest rates, we need to make use of a model which takes these into account when hedging long-dated options.

Table 1: The mean and standard deviation (in brackets) for daily delta hedging profit and loss for a 20-year ATM call option. The values are given for different hedging model choices (columns) and different world scenarios (rows).

		Hedging model		
		BS	Heston	HHW1
World	BS	0.0000 (0.0016)	-0.0916 (0.9063)	-0.0838 (0.9029)
	Heston	0.7480 (0.3662)	-0.0263 (1.2245)	-0.0111 (1.2198)
	HHW	0.6715 (2.2990)	-0.0743 (2.4935)	-0.0203 (1.8766)

Although there appears to be little difference between the Heston and HHW1 hedge under a BS and Heston world, the HHW1 hedge is slightly better. Comparison of the second and third columns of Table 1 (Heston and HHW1 hedging model choices) shows that the HHW1 model has a lower standard deviation than the Heston model in all three cases. The HHW1 model also has a mean closer to zero than the Heston model in all three

worlds. From this we can conclude that the inclusion of stochastic interest rates into a hedging model is not detrimental and can reduce the variance of the hedging PnL, as in the case of the HHW world.

The delta-vega hedging results are similar to the delta hedging results. Apart from the result that the BS hedge in a BS world performed poorly, the Heston and HHW1 hedges were better at reducing the hedge PnL in the BS world when hedging an ATM 20-year call option.

It is very important to note that the choice of the hedging instrument can have a very large effect on the performance of the hedge. We have chosen an ATM 5-year call option as the instrument to be used for the gamma and vega hedging. This was the longest dated instrument that we could realistically expect to find in the market.

6.2 Different hedging methods

To compare the different hedging method, we compare BS delta, delta-gamma, delta-vega hedging under a BS world. Similarly, we will also consider Heston hedging under a Heston world and HHW1 hedging under a HHW world, as presented in Figure 9. This allows us to compare the different hedging methods under each model. We will use the same instruments as before.

The first row of Figure 9 shows the hedging PnL for the different hedges in a BS world. Delta hedging performs the best in the BS world.[25] Delta-vega hedging in this world produces poor results. This is most likely due to the 5-year option being a poor volatility hedge for the 20-year option. If we need to hedge a long-dated option in a world where interest rates and volatility are constant, then delta hedging with a BS model would be the best hedging method to use.

When we move to a world containing stochastic volatility and constant

[25]This can be determined from Figure 9 as well as the standard deviations in Table 2.

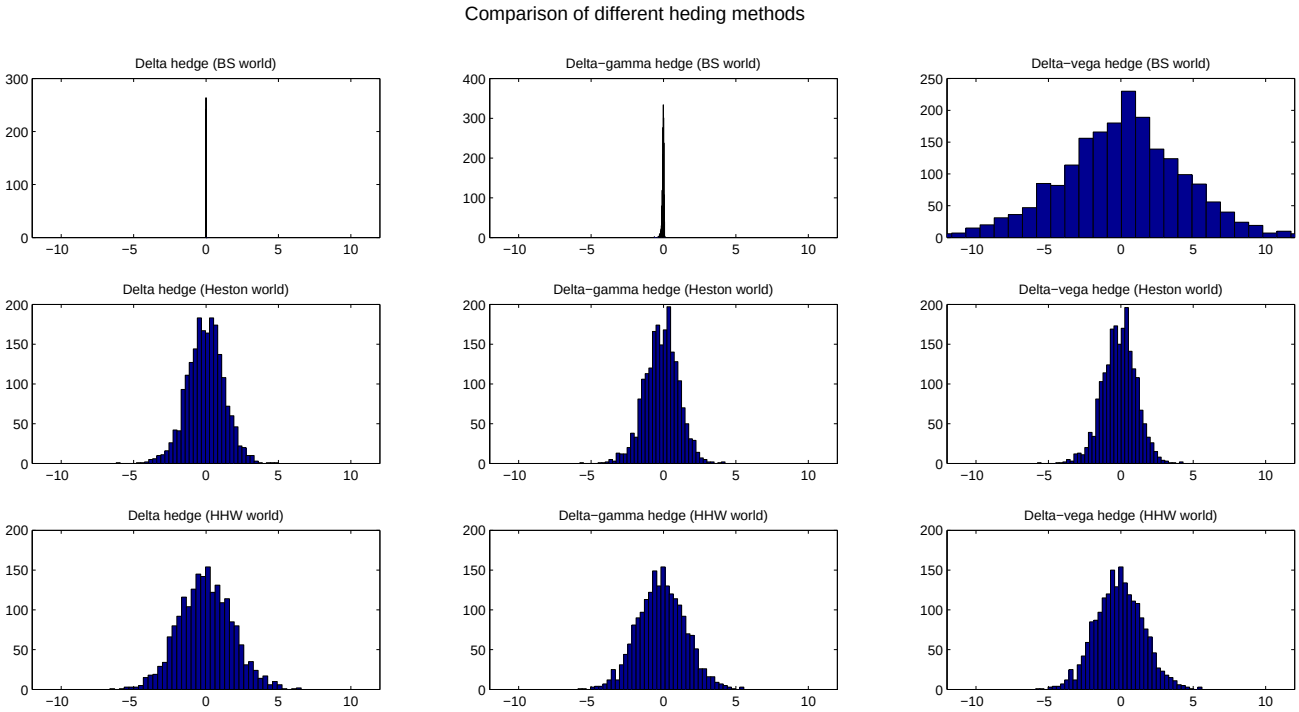


Figure 9: A comparison of the hedging profit and loss (PnL) for a 20-year ATM call option. The PnL is given for different hedging methods (columns) and different world scenarios (rows). In the case of delta-gamma and delta-vega hedging, the hedge instrument is a 5-year ATM call option. The model used for hedging is the same as the world model.

interest rates, it appears that all three hedging methods have similar results in Figure 9. If, however, we look at Table 2, we can see the differences: the delta hedge has the highest standard deviation in the Heston world (1.2732). The delta-gamma hedge has a slightly lower standard deviation than the delta-vega hedge (1.1636 versus 1.1637) (as shown in Table 2), but the mean of the delta-vega hedge (-0.1017) is closer to zero than the mean of the delta-gamma hedge (-0.1301).

In this world, we could make an argument for either delta-gamma or delta-vega hedging. There is a close link between gamma and vega, in that one cannot hedge both at the same time. To see this link, consider that gamma is the rate of change of the stock price and vega is the change in the

Table 2: The mean and standard deviation (in brackets) for daily hedging profit and loss for a 20-year ATM call option. The values are given for different hedging methods (columns) and different world scenarios (rows). In the case of delta-gamma and delta-vega hedging, the hedge instrument is a 5-year call option. The model used for hedging is the same as the world model.

		Hedge method		
		Delta	Delta-gamma	Delta-vega
World	BS	0.0000 (0.0016)	-0.0323 (0.0662)	0.0084 (4.5747)
	Heston	-0.0357 (1.2732)	-0.1301 (1.1636)	-0.1017 (1.1637)
	HHW	-0.0195 (1.8479)	-0.1353 (1.6315)	-0.1027 (1.6342)

volatility, which is the variability of the stock price. If there is a high gamma, then the stock price changes often, this is high volatility. This is due to the underlying algebraic relationship between the two Greeks. Although hedging either gamma or vega should protect us from the rate of the movements in the stock price, a choice between the two is needed. Kienitz and Wetterau (2012) state that vega hedging should be used in a world or model where stochastic volatility is present.

We now look at the HHW1 model in the third row of Figure 9. Adding stochastic interest rates to the world produces hedging PnL results that are very similar to those of the Heston model, but the differences in the hedging methods are slightly more pronounced. The delta hedge has a much wider distribution and is clearly the worst of the three methods. The delta-gamma and delta-vega hedges look similar once again. However, Table 2 shows that the delta-gamma hedge has a lower standard deviation (1.6315 versus 1.6342), but a mean further away from zero (-0.1353 versus -0.1027).

Once again, an argument could be made for either method, but we choose delta-vega hedging as it is the stochastic volatility that we would like to hedge. It is important to note that these results are for long-dated (20-year) call options and are very different when looking at shorter dated options.

6.3 Back-testing

To test the hedging performance of the model with real world data, we hedge a 20-year ATM call option on the JSE Top40 index. The mean value of this option at the start of the hedge (time zero) is R20557.44. The hedge instrument, where applicable, is a 5-year ATM call option, which has a mean value (at time zero) of R11064.68

To conduct the back-tests, we take observed Top40, v_0 and r_0 prices and values for 252 days and examine the hedging performance over the period. To get a large enough sample, we use overlapping 252-day periods: the first from 2 August 2010 to 19 September 2011, the second from 3 August 2010 to 20 September 2011, and so on. From 2 August 2010 to 31 May 2013, we can form 490 overlapping 252-day periods on which we run the back-tests. We perform two different back-tests: one with constant model parameters and the other with daily recalibration.

6.3.1 Model comparison

We first look at daily delta, delta-gamma and delta-vega hedging, where the model parameters are calibrated from the real world as in Section 5. We use the calibration results from Section 5.4 as the parameters for the hedging model; however, we hold these constant during the hedge.[26] This is done to examine whether or not the previous simulated hedging results hold up under real world model parameters.

Histograms of the hedging PnL for this back-test are presented in Figure 10. The PnL under BS and Heston hedges are provided for completeness, however, we do not discuss these results in detail. It is worth noting that the HHW1 model outperforms both the BS and the Heston model under all

[26] We do not recalibrate for every day of the hedge. On day one, we take the calibration results for this day and then back-test daily hedging for 252 days. We then take calibration results from day two and back-test hedging for 252 days.

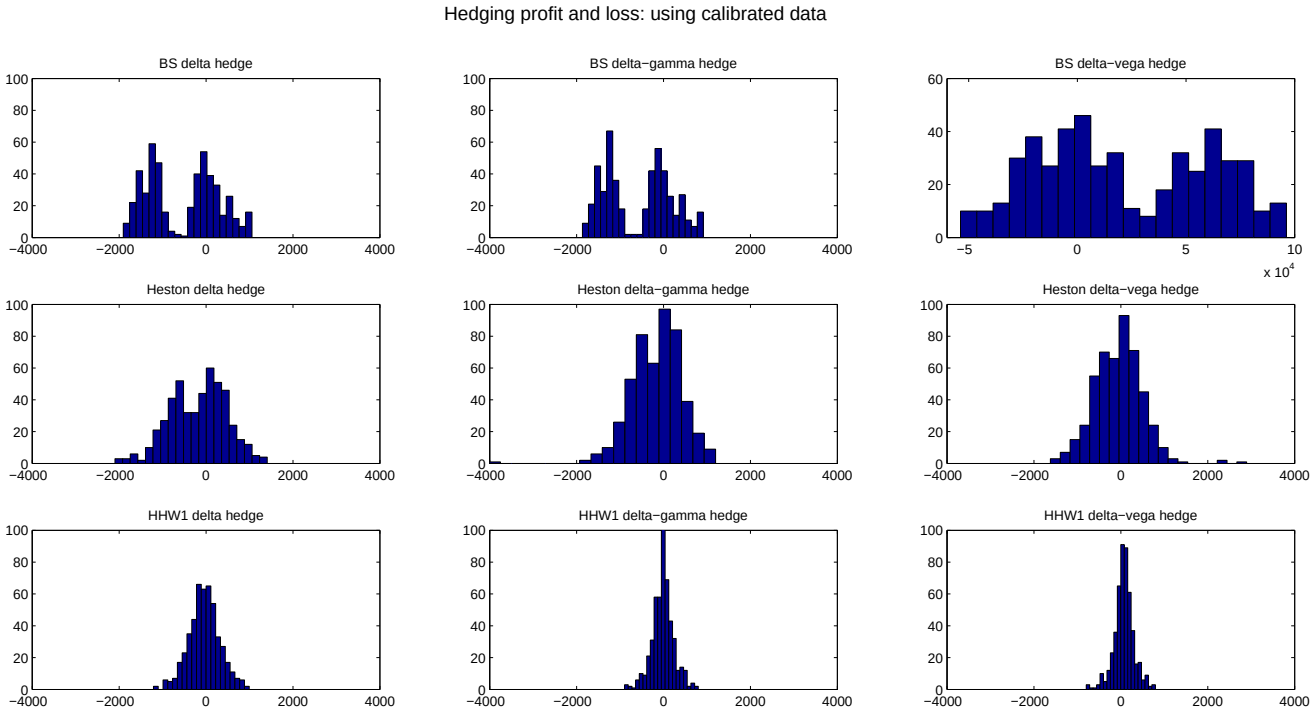


Figure 10: Back-tested daily hedging profit and loss (PnL) for a 20-year ATM call option on the JSE Top40. PnL is given for different hedging methods (columns) and different hedging models (rows). Please take note of the different scale used for the BS delta-vega hedge.

three hedging methods. This can also be observed in Table 3.

We now focus on the HHW1 model results. As expected, the delta hedge is worse than the second order hedges. The delta hedge has a much wider distribution and Table 3 confirms it has the highest standard deviation (356.660). The delta-vega hedge has the lowest standard deviation of the three hedging methods (225.850 versus 356.660 and 251.540). Interestingly, the delta-vega hedge had a higher standard deviation in the simulations from Section 6.2, but as Kienitz and Wetterau (2012) recommended, the delta-vega hedge performs better in practice. The mean of the delta-vega hedge is, however, further away from zero than the delta or delta-gamma mean (73.2518 versus -36.4037 and -11.3955 respectively).

Although it may appear, when looking at the scale of the PnL, that the

Table 3: The mean and standard deviation (in brackets) of the back-tested hedging profit and loss (PnL) for a 20-year ATM call option on the JSE Top40. The values are given for different hedging methods (columns) and different hedging models (rows).

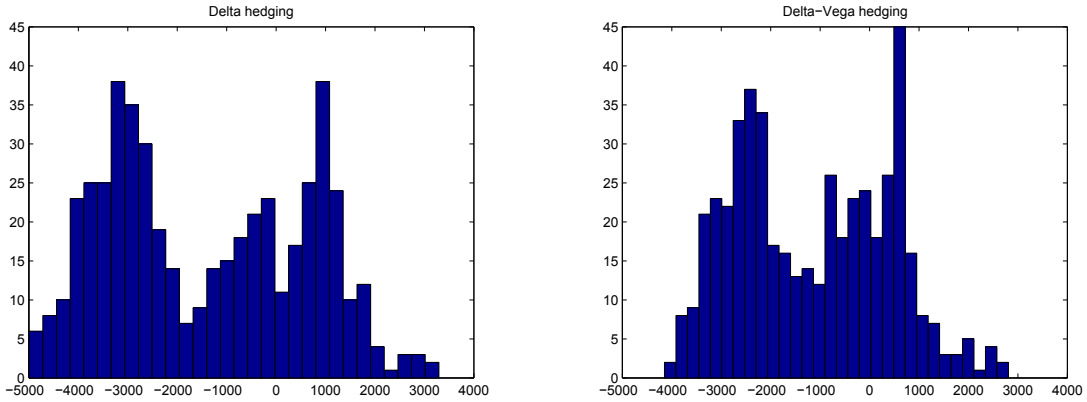
	Hedge method		
Model	Delta	Delta-gamma	Delta-vega
BS	-529.201 (805.851)	-579.711 (757.088)	21312.6 (39227.8)
Heston	-174.523 (649.161)	-160.166 (578.815)	-56.684 (553.409)
HHW	-36.404 (356.660)	-11.396 (251.540)	73.252 (225.850)

hedges in the third row of Figure 10 are worse than those from the third row Figure 9 in Section 6.2, this is not the case. In Section 6.2, the normalised standard deviation (standard deviation per unit of option value at time zero) for the HHW1 delta-vega hedge is 0.0246. The normalised standard deviation for the delta-vega hedge in this back-test is 0.0110, which is much lower. We therefore actually have a better hedge when using the calibrated parameters.

6.3.2 Daily recalibration

To examine the model performance under daily recalibration, we run daily delta and delta-vega back-tests with the model parameters being recalibrated each day. The PnL histograms for these tests are given in Figure 11.

These hedging results do not look approximately normally distributed. This is most likely due to the simple calibration technique that was used. As explained in Section 5.4, the stability of the model parameters is important for hedging and, although the model has periods of stability, there are many large jumps in the parameters. These jumps could be causing large losses and hence the second peak (the peak centered around -3000) of the



(a) Delta (b) Delta-vega

Figure 11: Back-tested profit and loss results of daily delta and delta-vega hedging with daily recalibration of a 20-year ATM call option on the JSE Top40.

histogram. If we assume that this is the cause of the second peak, we can assume that, if we had stable model parameters, this second peak would not be here. Looking at Figure 11, we can see that the peak centered around zero looks very similar to the previous hedging results.

Table 4: The mean and standard deviation of the back-tested hedging profit and loss (PnL) for a 20-year ATM call option on the JSE Top40. Daily recalibration was performed during the hedge.

	Hedge method	
	Delta	Delta-vega
Mean	-1399.78	-1158.36
Standard deviation	1999.13	1553.63

The delta-vega hedge outperforms the delta hedge as expected. It has a lower standard deviation (1553.63 versus 1999.13) and a mean closer to zero (-1158.36 versus -1399.78), as shown in Table 4. The delta-vega hedge has a normalised standard deviation of 0.0756, which is much higher than in all previous tests. Although these hedges perform poorly, they still confirm the

observations made earlier. Comparing this back-test with the first back-test, we draw the conclusion that, given constant or relatively stable parameters, the HHW1 model outperforms the BS and Heston models.

One must remember that these back-tests are only one possible realisation of the world and cannot be directly compared to the results from earlier Sections. Nonetheless, these back-testing results are promising and show that the model would have been able to hedge the option quite well over the past three years: better, in fact, than if the Heston model or Black-Scholes models were used.

7 Summary and Conclusions

The approximations made by Grzelak and Oosterlee (2011) allow the combination of a stochastic volatility model and a stochastic interest rate model, in an affine form, with non-zero correlation between the equity and interest rate processes. They present the ChF for the HHW1, HHW2 and HHW3 models in their paper.

We implement efficient pricing of contingent claims under the HHW1 and HHW3 models with the Fourier-Cosine method of option pricing. We do not implement the HHW2 model as it involves numerically solving a system of ODEs numerous times, and although this can be done, it is currently too slow for our purposes. More research into these ODEs is recommended, as, if an approximation or general solution could be found, then the HHW2 model could be just as fast as the other models.

After making some slight changes to the ChF, we implemented it in MATLAB. We did not use the $\widetilde{\Lambda}_t$ approximation, and instead evaluated the integral in the ChF for the HHW models. This allows us to use the models even if the Feller condition does not hold, which happens often in practice.

The HHW3 model is 15-20 times slower than the HHW1 model, as we were not able to vectorise the HHW3 implementation. We could therefore not make use of MATLAB's optimised vector calculations. It takes 0.122 seconds on average to price a single option under the HHW1 model. Using parallelisation, it takes 6 seconds to price 10000 options with the HHW1 model. Thus, we have a sufficiently efficient implementation of the HHW model.

Compared to a full scale Monte Carlo implementation, the HHW1 and HHW3 models perform accurately. Both models price options whose value is very close to the Monte Carlo value and deviates, as expected, from the Black-Scholes value.

We use a very simple calibration technique and conclude that it is pos-

sible to get stable model parameters from a HHW calibration. We also note that both the HHW1 and HHW3 models are efficient enough to be used for daily calibration.

The HHW models outperformed both the Black-Scholes and the Heston model when conducting daily delta hedging. This holds true for a simulated world, as well as during back-testing. Even when interest rates were constant (i.e. in a Heston world), the HHW model outperformed the Heston model when hedging a long-dated (20-year) call option. When comparing different hedging methods, we confirmed the idea that vega hedging performs better than gamma hedging in a world containing stochastic volatility (Kienitz and Wetterau, 2012). This is especially true when dealing with long-dated options.

Back-tests confirmed the simulated hedging results and in most cases actually produced better results. A daily recalibration back-test, which mimics what would happen in practice, confirms the previous hedging observations. Even though this back-test produced relatively poor hedging results, due to the simple calibration procedure, it was very insightful. It confirmed that delta-vega hedging is the best hedging method, and that the model is fast enough to price under the constraints of daily recalibration. The hedge results, although much worse than all other tests, were not terrible, and, if we ignore the second peak, look very similar to all previous hedging tests. From this, we conclude that, given stable parameters from a better calibration technique, it is reasonable to assume that the HHW models will outperform the Heston and Black-Scholes models when hedging a long-dated (20-year) call option.

Overall, the HHW model has performed well. Implementation involved a few tricks and adjustments, but after introducing some parallelisation, the model prices options efficiently. The HHW model has promising parallelisation results, as well as the ability to be further accelerated by advanced

parallelisation methods. As well as fast, the model is also accurate when prices are compared to Monte Carlo simulated prices. Therefore, we have achieved our aim of implementing an efficient and accurate HHW model.

Calibration of the HHW model should be relatively simple, as the Hull & White component is calibrated first and then the Heston component is calibrated. It is recommended that calibration is explored prior to implementing the HHW model in practice. Finally, the hedging results are very promising, as the model outperformed the Heston and Black-Scholes models in numerous tests.

In conclusion, the aims of this dissertation have been achieved. We recommend the implementation of the HHW model over the Heston and Black-Scholes models when pricing and hedging long-dated vanilla European options.

References

Bakshi, G., Cao, C. and Chen, Z. (2000). Pricing and hedging long-term options, *Journal of Econometrics* **94**(1-2): 277–318.

Black, F. (1976). The pricing of commodity contracts, *Journal of financial economics* **3**(1): 167–179.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities, *Journal of Political Economy* **81**(3): pp. 637–654.
URL: *http://www.jstor.org/stable/1831029*

Brigo, D. and Mercurio, F. (2007). *Interest rate models-theory and practice: with smile, inflation and credit*, second edn, Springer.

Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985). A theory of the term structure of interest rates, *Econometrica* **53**(2): pp. 385–407.
URL: *http://www.jstor.org/stable/1911242*

Fang and Oosterlee (2008). A novel pricing method for european options based on fourier-cosine series expansions, *MPRA Paper 7700*, University Library of Munich, Germany.
URL: *http://ideas.repec.org/p/pra/mprapa/7700.html*

Feller, W. (1951). Two singular diffusion problems, *The Annals of Mathematics* **54**(1): 173–182.

Grzelak, L. A. and Oosterlee, C. W. (2011). On the heston model with stochastic interest rates, *SIAM Journal on Financial Mathematics* **2**(1): 255–286.

Grzelak, L. A., Oosterlee, C. W. and Weeren, S. V. (2012). Extension of stochastic volatility equity models with the Hull&White interest rate process, *Quantitative Finance* **12**(1): 89–105.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options, *The Review of Financial Studies* **6**(2): 327–343.
URL: *http://www.jstor.org/stable/2962057*

Hull, J. and White, A. (1990). Pricing interest-rate-derivative securities, *The Review of Financial Studies* **3**(4): pp. 573–592.
URL: *http://www.jstor.org/stable/2962116*

Hunter, C. (2005). Hybrid derivatives, *The Euromoney Derivatives and Risk Management Handbook*.

Jackel, P. (2004). Stochastic volatility models: Past, present and future, *The Best of Wilmott I: Incorporating the Quantitative Finance Review* **1**: pp. 379–390.

Jamshidian, F. (1989). An exact bond option formula, *The journal of Finance* **44**(1): 205–209.

Kienitz, J. and Wetterau, D. (2012). *Financial Modelling: Theory, Implementation and Practice with MATLAB Source*, Wiley Finance.

Lord, R., Koekkoek, R. and Dijk, D. V. (2010). A comparison of biased simulation schemes for stochastic volatility models, *Quantitative Finance* **10**(2): 177–194.
URL: *http://ssrn.com/abstract=903116*

MATLAB (2013). *version 8.1.0.604 (R2013a)*, The MathWorks Inc., Natick, Massachusetts.

McWalter, T. A. (2013). Numerical Methods II, *Lecture notes*, University of Cape Town, South Africa.

Oehlert, G. W. (1992). A note on the delta method, *The American Statistician* **46**(1): pp. 27–29.
URL: *http://www.jstor.org/stable/2684406*

OPTI-NUM solutions (Pty) Ltd (2013). Parallel Computing with MATLAB for Finance, *Private seminar*.
URL: *http://www.optinum.co.za/*

Wang, G. (2011). *An Equity and Foreign Exchange Heston-Hull-White model for Variable Annuities*, unpublished PhD thesis, Delft Institute of Applied Mathematics.

Zhu, J. (2000). Modular pricing of options, *Wirtschaftswissenschaftliches Seminar*.

Appendices

A Some mathematical deffinitions

In this appendix we will define a characteristic function (ChF) and what it means for a model to be affine.

A.1 Characteristic Function

If X is a random variable then its characteristic function $\phi_X : \mathbb{R} \to \mathbb{C}$ is defined as:

$$\phi_X(u) := \mathbf{E}\left(e^{iuX}\right), \quad u \in \mathbb{R},$$

where $\mathbf{E}$ denotes the expectation, i is the imaginary unit and u is the ChF's argument.

The probability density function of any random variable, $X \in \mathbb{R}$, is completely described by the ChF of that random variable. This allows us to use the ChF in place of the probability density function when evaluating the price of an option

A.2 Affinity

For affinity to hold, the expectations and the covariance matrix of the system need to be linear in the state space variables (the variables that the model describes). If a model is affine its ChF exists and is log-linear, i.e. is of the form:

$$\phi(u, \tau, \underline{s_t}) = \exp\left[\mathcal{A}(u, \tau) + \underline{\mathcal{B}}(u, \tau) \cdot \underline{s_t}\right],$$

where, $\tau := T - t$ is the time to maturity, $\underline{s_t}$ is the state space vector, $\mathcal{A}(u, \tau)$ and $\underline{\mathcal{B}}(u, \tau)$ are the coefficient functions, and $\underline{x}$ denotes that x is a vector. If we can make a model affine we know the form of the corresponding ChF, this gives us an approach to finding the solution of the ChF.

B Tables of call option prices

Table 5 and Table 6 present call option prices under the HHW1 and HHW3 models, as well as under a Monte Carlo (MC) implementation.

Note that the differences between the HHW1 and MC prices are almost always within three standard deviations of the MC prices.

Table 5: Prices of a $T = 1$ year call option priced with the COS method under HHW1 and HHW3 and a Monte Carlo scheme. The standard deviations of the Monte Carlo prices are shown in brackets. The differences between the HHW1 prices and the other two prices are also given. Model parameters: $S_0 = 100$; $v_0 = 0.0175$; $r_0 = 0.07$; $\kappa = 1.5768$; $\bar{v} = 0.0398$; $\gamma = 0.0571$; $\lambda = 0.05$; $\theta = 0.07$; $\eta = 0.005$; $\rho_{x,v}$ = -0.5711; $\rho_{x,r} = 0.2$; $\rho_{r,v} = 0.3$.

	Price			Difference	
Strike	HHW1	HHW3	MC	HHW3	MC
50	53.3802	53.3802	53.3858 (0.0076)	0.0000	0.0056
55	48.7188	48.7188	48.6953 (0.0075)	0.0000	-0.0235
60	44.0594	44.0594	44.0575 (0.0076)	0.0000	-0.0019
65	39.4076	39.4076	39.3795 (0.0077)	0.0000	-0.0281
70	34.7773	34.7772	34.7742 (0.0085)	-0.0001	-0.0030
75	30.1978	30.1974	30.2189 (0.0097)	-0.0004	0.0215
80	25.7199	25.7193	25.7074 (0.0116)	-0.0006	-0.0119
85	21.4184	21.4175	21.3814 (0.0139)	-0.0009	-0.0361
90	17.3856	17.3847	17.3550 (0.0175)	-0.0009	-0.0297
95	13.7185	13.7175	13.6469 (0.0207)	-0.0010	-0.0706
100	10.4998	10.4991	10.4725 (0.0239)	-0.0007	-0.0266
105	7.7828	7.7825	7.7198 (0.0256)	-0.0003	-0.0627
110	5.5814	5.5816	5.5510 (0.0251)	0.0002	-0.0306
115	3.8711	3.8717	3.8524 (0.0226)	0.0006	-0.0193
120	2.5968	2.5978	2.5806 (0.0192)	0.0010	-0.0172
125	1.6856	1.6868	1.6894 (0.0160)	0.0012	0.0026
130	1.0597	1.0609	1.0464 (0.0126)	0.0012	-0.0145
135	0.6458	0.6469	0.6311 (0.0099)	0.0011	-0.0158
140	0.3820	0.3830	0.3857 (0.0076)	0.0010	0.0027
145	0.2196	0.2204	0.2292 (0.0059)	0.0008	0.0088
150	0.1229	0.1234	0.1325 (0.0044)	0.0005	0.0091

Table 6: Prices of a $T = 10$ year call option priced with the COS method under HHW1 and HHW3 and a Monte Carlo scheme. The standard deviations of the Monte Carlo prices are shown in brackets. The differences between the HHW1 prices and the other two prices are also given. Model parameters: $S_0 = 100$; $v_0 = 0.0175$; $r_0 = 0.07$; $\kappa = 1.5768$; $\bar{v} = 0.0398$; $\gamma = 0.0571$; $\lambda = 0.05$; $\theta = 0.07$; $\eta = 0.005$; $\rho_{x,v}$ = -0.5711; $\rho_{x,r} = 0.2$; $\rho_{r,v} = 0.3$.

	Price			Difference	
Strike	HHW1	HHW3	MC	HHW3	MC
50	75.2871	75.2847	75.1201 (0.1156)	-0.0024	-0.1646
55	72.8989	72.8957	72.7099 (0.1180)	-0.0032	-0.1858
60	70.5437	70.5396	70.3364 (0.1184)	-0.0041	-0.2032
65	68.2258	68.2208	68.3918 (0.1241)	-0.0050	0.1710
70	65.9492	65.9433	66.0793 (0.1246)	-0.0059	0.1360
75	63.7175	63.7106	63.8027 (0.1266)	-0.0069	0.0921
80	61.5335	61.5257	61.6586 (0.1294)	-0.0078	0.1329
85	59.3999	59.3912	59.1382 (0.1271)	-0.0087	-0.2530
90	57.3186	57.3090	57.5263 (0.1304)	-0.0096	0.2173
95	55.2912	55.2809	54.9223 (0.1293)	-0.0103	-0.3586
100	53.3190	53.3080	52.9550 (0.1316)	-0.0110	-0.3530
105	51.4027	51.3912	51.5674 (0.1373)	-0.0115	0.1762
110	49.5429	49.5309	49.5189 (0.1373)	-0.0120	-0.0120
115	47.7396	47.7272	47.7603 (0.1402)	-0.0124	0.0331
120	45.9928	45.9801	45.7863 (0.1407)	-0.0127	-0.1938
125	44.3021	44.2893	44.3100 (0.1440)	-0.0128	0.0207
130	42.6670	42.6541	42.7661 (0.1452)	-0.0129	0.1120
135	41.0868	41.0739	40.9534 (0.1438)	-0.0129	-0.1205
140	39.5605	39.5478	39.2994 (0.1443)	-0.0127	-0.2484
145	38.0873	38.0747	37.8964 (0.1474)	-0.0126	-0.1783
150	36.6660	36.6537	36.1094 (0.1455)	-0.0123	-0.5443

Acknowledgements

I would like to thank my supervisors, Moses dos Santos and Marchand van Rooyen, for their input and advice. Moses dos Santos, from Old Mutual, deserves further thanks for supplying vital data. Special thanks goes to Tom McWalter, for the discussions we had and various insightful points he added.

I am very grateful to Sarah Phelp, for her continuous support and for a complete revision of the draft, and to Ralph Rudd, for facilitating the printing and sumbission of the final copy of this dissertation. Finally, I would like to thank Jessica Micklem for playing her part in keeping me sane throughout the duration of this project.

Declaration

1. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.

2. Each significant contribution to, and quotation in, this dissertation from the work of other people has been cited and referenced.

3. This dissertation is my own work.

4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

5. This dissertation is being submitted for the Degree of Master of Philosophy in the University of Cape Town, South Africa.

6. It has not been submitted before for any degree or examination in any other University.

Name: Sheldon Maze
Student number: MZXSHE001
Signature: _____
Date: 21st May 2014